FINAL REPORT

Development of an Index of Biotic Integrity for Fish

Assemblages for Wadeable Streams in Arkansas' Ozark Highlands Ecoregion

Prepared for

Tetra Tech, Inc.

For the subcontract for EPA Contract #68-C98-111 WA#1-10

by

Edmund J. Pert[1] and Daniel C. Dauwalter[2]
University of Arkansas at Pine Bluff
Department of Aquaculture/Fisheries

11-1-01

Current Contact Information for Authors:

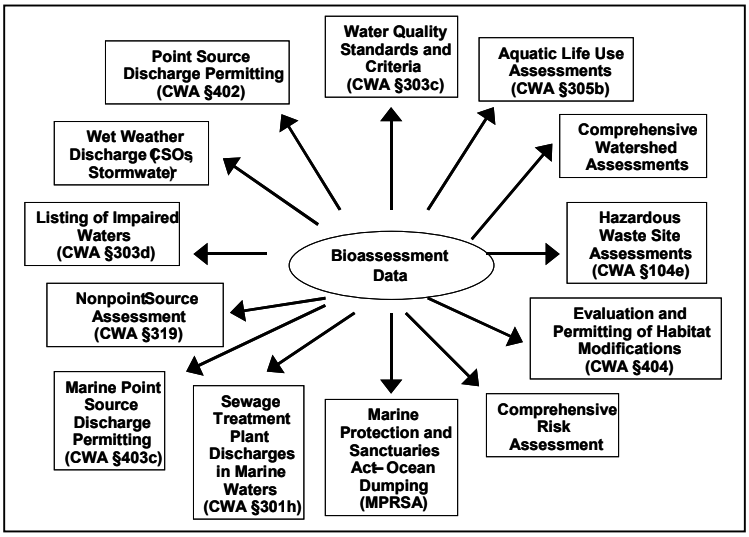
[1]Edmund J. Pert
California Department of Fish and Game
1416 Ninth Street, 12th floor
Sacramento, CA 95814
Epert@dfg.ca.gov
Phone (916) 651-8689

[2]Daniel C. Dauwalter
University of Arkansas at Pine Bluff
Department of Aquaculture/Fisheries
1200 N. University Dr.
Pine Bluff, AR 71601
ddauwalter@uaex.edu
Phone (870) 543-8107

Table of Contents

Executive Summary

Monitoring of biological communities integrates effects of different pollutant stressors such as excess nutrients, toxic chemicals, increased temperature, and excessive sediment loading. Thus, biomonitoring provides an overall measure of the aggregate impact of those stressors. Biological communities respond to stresses of all degrees over time and, therefore, offer information on perturbations not always obtained with episodic water chemical measurements or discrete toxicity tests. The central purpose of assessing the biological condition of aquatic communities is to determine how well a water body supports aquatic life.



Use of Bioassessment in State Water Quality Programs

Biological communities reflect overall ecological integrity (i.e., chemical, physical, and biological integrity). Therefore, bioassessment results directly assess the status of a waterbody relative to the primary goal of the Clean Water Act (CWA). Biological assessments are crucial to evaluating ecosystem health and provide crucial water quality planning information for managing more complex water quality problems (see graphic listing water quality programs).

In Arkansas, bioassessment has been used for more than two decades enabling the Department of Environmental Quality (ADEQ) to monitor and assess the state's waters. However, a full development of an Index of Biotic Integrity (IBI) has not been completed to date. The value of a calibrated IBI is that the index can be used by the ADEQ to assess the biological condition of its regional streams for which the IBI was calibrated. The IBI may also be used to prioritize regional waters for protection, maintenance, and/or restoration. This report documents the process of developing an IBI for fish assemblages for use in Arkansas' Ozark Highlands Ecoregion. The first Chapter focuses on the development of standardized sampling protocols, and the second Chapter discusses the development and application of the IBI.

The highlights of this developmental biological research project are as follows:

- Results indicate that when sampling in wadeable Ozark Highland streams via backpack electrofishing a distance of 51 mean stream widths (MSW), i.e., a

length of 51 times the average wetted width of the stream, will enable a collection of approximately 95% of the species in, and nearly perfectly reflect the fish-assemblage structure of that stream site. An equivalent effort should be sampled in larger streams that require barge electrofishing.

A distance of 51 MSW is equivalent to 16.1 mean bankfull widths (bankfull width measured as the upper bank level of the stream), or 5.7 riffle-pool sequences.

This research helped to establish the framework for standardized sampling protocols to be used in Arkansas in support of the IBI.

Fish-collection data from 96 stream sites in the Ozark Highlands Ecoregion were partitioned to both develop and validate the IBI.

Ten out of 39 potential biological metrics were found to be the most informative, sensitive, and robust for the IBI. Trophic metrics contributed the most information to IBI scores.

Core Metrics of the IBI

Species Richness and Composition

Total species
Total darter, sculpin, & madtom species
% individuals as green sunfish, bluegill, yellow bullhead, & channel catfish
% individuals as stonerollers
% individuals as intolerant species

Reproductive Composition

Total mineral spawning species

Trophic Composition

% individuals as top carnivores
% individuals as algivorous/herbivorous, invertivorous, & piscivorous
% individuals as invertivores

Fish Abundance and Condition

% individuals with black spot or anomaly

All metrics were transformed from absolute values to unitless scores ranging from 0 to 10; where, 0 indicates a strong deviation from reference condition, and 10 represents best attainable condition.

The IBI was an aggregate of 10 core metrics and was calculated to score from 0 to 100 (the sum of the maximum score of the metrics); where, 0 indicates no fish and 100 represents best attainable condition.

- The range in IBI scores for the population of reference sites was 79 to 100. Using a population statistic of the 25th percentile and distribution of reference-site IBI scores, the threshold IBI score for differentiating impairment from unimpairment is proposed to be 85. A scoring range from 83-87 is recommended for professional judgement of stream impairment designations.

- A range of 4 IBI scoring units is recommended for professional judgement of site quality when IBI scores fall at or near a qualitative classification threshold. This range was determined by assessing the within-site precision of the IBI, and should be centered on each threshold score.

- A sampling distance deviating from 51 MSWs, or an equivalent effort, can potentially affect IBI scores and qualitative site classifications.

- Nutrients, land use, road density, and sedimentation were most consistently correlated with the IBI metrics.

This bioassessment framework, i.e., the Ozark Highlands IBI, has been tested on the streams in this ecoregion, and is ready to be validated and implemented in other parts of Arkansas. The IBI is based on EPA procedures and is a cost-effective biomonitoring tool designed to enable the state of Arkansas to better assess and monitor stream quality throughout Arkansas. The future of the Ozark Highlands IBI is the broader development and refinement for statewide implementation. The link between biological indicators and water regulation is integral to water resource protection.

CHAPTER 1

Relations Between Sampling Effort and Fish Species Richness and Assemblage Structure in Wadeable, Ozark Highland Streams of Arkansas

Abstract

Standardized sampling protocols are essential for consistent characterization of stream fish assemblages, especially when making spatial or temporal comparisons. We sampled 15 stream sites in the Ozark Highlands ecoregion of Arkansas and examined the effect of increased sampling effort on estimates of fish species richness and assemblage structure, two measures often used when characterizing fish assemblages. For each habitat unit (i.e., riffle or pool) we measured the habitat unit length, mean stream width (MSW), and bankfull width at each stream site. Each site was 75 MSWs in length and was divided into 15 consecutive segments that were five MSWs long. We collected fish in each segment by using a pulsed, DC backpack electroshocker. For each site the percent of total species collected and percent community similarity to the entire fish assemblage were calculated with the addition of each consecutive segment in both upstream and downstream directions. We determined the number of species, and which taxa, were most likely to be missed when sampling the average distance needed to characterize regional fish assemblages. We also compared our field MSW estimates to a habitat-based length-weighted MSW and assessed the percentage error in our MSW estimates. When using our field MSW estimates a distance of 51.0 MSWs will, on average, collect 95% of the species in a stream site. When using the length-weighted MSWs, 46.1 MSWs need sampling. Both MSW estimates are equivalent to 16.1 mean bankfull widths or 5.7 riffle-pool sequences. Obtaining 95% community similarity requires sampling a length equivalent to 12.1 - 13.4 MSWs, 4.2 bankfull widths, or 1.5 riffle-pool sequences. Our data indicated that, when assuming 100% of species were sampled in all 15 segments (i.e, 67.7 – 75.0 MSWs), an average of 1.0 (SD = 1.1) species

would be missed when sampling 46.1 – 51.0 MSWs. Although the true MSW remains unknown, based on the sample size and steam width variability of our samples, we determined with 95% confidence that our field MSW estimates were within 13.3% of the true population mean, while our length-weighted MSW estimates were within 8.0% of the true population mean.

Introduction

The species-area relation (Arrhenius 1921; Gleason 1922) is an important consideration for characterizing stream fish assemblages (Angermeier and Schlosser 1989). Consequently, it is important to use standardized sampling efforts to assess temporal and spatial fish-assemblage differences among comparable stream sites. This is because sampling effort affects the number of species collected, as well as other assemblage characteristics (Angermeier and Schlosser 1989; Lyons 1992; Paller 1995; Angermeier and Smogor 1995; Patton et al. 2000).

Characterizing fish assemblages requires consistent, and preferably representative, sampling in order to minimize potential biases when comparing stream sites. Using consistent sampling protocols is imperative for unbiased comparisons because of the increased likelihood of collecting more species with increasing sampling effort, or missing species by decreasing sampling effort. Because most natural resource agencies lack the resources to sample any more than is absolutely necessary, efforts to establish rapid sampling protocols that are consistent and representative have increased in the area of biological monitoring (e.g., Plafkin et al. (1989)). However, development of standardized sampling protocols for fish assemblages has lagged behind other taxa such as benthic macroinvertebrates. Despite a lack of concerted efforts to develop standardized sampling protocols for fish assemblages in streams, some studies have contributed to the development of standardized sampling protocols in wadeable streams by studying the effects of sampling effort on species richness and relative abundance in fish assemblages. Meador et al. (1993) summarized the work to date that contributed to standardizing sampling efforts in wadeable streams; they recommended sampling a

minimum stream length of 150 m and a maximum length of 300 m in wadeable streams for the National Water-Quality Assessment Program (NAWQA). They noted that sampling distance will depend on individual stream characteristics and should not exceed a distance at which crew fatigue will affect sampling efficiency. Although these recommended sampling lengths were based on literature reports, they were somewhat arbitrary because there were no studies that could reliably address the issue of a standard sampling effort across a broad spatial scale. Additional studies that tested the effect of sampling effort on certain fish-assemblage characteristics (e.g., species richness and relative abundance) indicated that differences in sampling effort will alter the fish assemblage characterization (Angermeier and Schlosser 1989; Lyons 1992; Paller 1995; Angermeier and Smogor 1995; Patton et al. 2000). Differences in sampling efforts can also affect stream-fish-community indices, such as the Index of Biotic Integrity (IBI) (Karr 1981; Angermeier and Karr 1986).

The goal of this study was to develop an approach to standardize sampling protocols for stream fishes in wadeable streams of Arkansas' Ozark Highlands ecoregion. We based our approach on typical collection techniques used by agencies in Arkansas, and we emphasized developing a protocol that would minimize the effort required to collect most of the species present at a site. Our objectives for this study were three-fold: 1) to determine the average sampling effort (i.e., relative stream distance) needed to collect 95% of the species present and obtain 95% similarity between sampled and representative fish assemblages in wadeable, Ozark Highland streams of Arkansas, 2) based on the patchiness of species' local distributions, determine the taxonomic groups most likely to be missed when using the average sampling effort needed to characterize

fish assemblages, and 3) determine, with 95% confidence, the accuracy of our initial field estimates of mean stream width (MSW), and determine the sample sizes needed to estimate MSW at different levels of precision around the true population mean.

Study Area

The Ozark Highlands ecoregion is located in north-central and north-western Arkansas (Figure 1.1). Mountainous terrain, steep gradients, and fractured limestone geology characterize (Robison and Buchanan 1988) and form the fast-flowing, spring-fed streams of the region. Land use is a major cause of regional water-quality problems. Many of these problems result from high animal production wastes (e.g., poultry) that have a potential to contaminate regional surface and ground waters because of local geology (ADPCE 1996). The ichthyofauna in this ecoregion is highly diverse, consisting of at least 89 species with cyprinids, percids, and centrarchids contributing most to Ozark streams' relative abundance (Giese et al. 1987).

Methods

Data Collection

We sampled 15 wadeable streams throughout the Ozark Highlands ecoregion in Arkansas (Figure 1.1). We selected sample sites to include streams that differed in size and anthropogenic disturbance levels, as well as were widely distributed across the ecoregion. We sampled these sites between 6 June and 1 August 2000.

We conducted ground reconnaissance to ensure that qualitative riparian condition and land use adjacent to our selected stream sites varied among sample sites. We also

quantified land use in each site's watershed with a Geographic Information System (GIS). We delineated each sample site's watershed by using ArcView® version 3.2 and the Spatial Analyst extension. We calculated the percent land use in each delineated watershed by overlaying the watershed and land-use layers (Table 1.1). We used the Arkansas Gap Analysis (AR-GAP) land-cover data layer (Smith et al. 1998; Weih 2001) for land-use determinations. Landsat Thematic Mapper (TM) data, mostly from 1992, was used as the base data layer to develop AR-GAP land-cover classifications. The final AR-GAP data layer includes 36 land cover classes with a 100 ha resolution. We grouped the AR-GAP land-cover classifications into agriculture, forested, and urban land-use classes.

We measured some physical features of each sample site in a stream reach of 75 MSWs. Based on literature reports (Lyons 1992; Paller 1995; Angermeier and Smogor 1995) and the Arkansas Department of Environmental Quality's (ADEQ) sampling protocol for this ecoregion, this stream length should be sufficient to obtain a representative sample of each fish assemblage when collected by backpack electrofishing. We classified habitat units as either riffles or pools in order to minimize classification bias (Roper and Scarnecchia 1995). Each habitat unit was classified as either a riffle or pool depending on adjacent habitat units, the relative water depths and velocities, and surface turbulence of all habitat units in a given stream. Based upon on-site discussions about habitat typing, this classification scheme resulted in consistent and repeatable classification of riffle-pool sequences. We also measured the total length of each habitat unit. At transects in each habitat unit we measured stream width (wetted width) and mean water depth according to Platts et al.'s (1983) stratified transect design.

The number of transects per habitat unit ranged from one to eight depending on habitat unit length and heterogeneity of channel width and depth. For example, homogenous habitat types required fewer transects for characterization than did heterogeneous habitat types (Platts et al. 1983). The same was true for short versus long habitat units. We attempted to keep distances between transects as evenly distributed as possible given differing habitat unit lengths in order to accurately represent the mean wetted width. We measured bankfull width at one transect in a habitat unit. We included bankfull width as an approximation for comparison with other studies. Bankfull width was estimated by observing bank vegetation, sediment particle size characteristics, and changes in bank slope. Despite our efforts to consistently identify bankfull width, we believe bankfull width was our least-repeatable measure of relative stream size (see Simonson et al. (1993) and Simonson et al. (1994)). Because of the uncertainty in estimating bankfull width, and the fact that bankfull width appeared to be relatively homogeneous within habitat units, we felt that measuring bankfull width more than once in a habitat unit would not improve the accuracy of our estimate.

We also measured discharge at one transect in each site. Each transect was located in a stream section with a relatively smooth-bottomed channel cross-section that was divided into 20-25 cells according to Gallagher and Stevenson (1999). Transect-section areas were determined and water velocity was measured using a Marsh-McBirney FLOW-MATE^{TM} 2000 portable flowmeter. Sectional discharges were determined measuring the depth, mean column velocity (0.6 of depth), and cell width, multiplying those values within each cell, and then cells were summed for total discharge.

After measuring habitat parameters we calculated the arithmetic mean of all wetted-width measurements. This mean was then used as a field estimate of MSW. Stream reaches of 75 MSWs were divided into 15 segments that were each five MSWs long. Stream segmentation always started at the upstream end of each site. Therefore, the final segment (i.e., segment 15) always ended at the top or bottom of a riffle. These riffles were assumed to limit fish movement in the final segment when sampling in an upstream direction.

In each segment, a four-person crew, operating a pulsed DC, Smith-Root backpack electroshocker outfitted with two anodes, collected fish by wading upstream. Two crew members carried dipnets to collect fish. Riffles were fished by placing two dipnets side-by-side while anodes were waved in an upstream to downstream direction towards the dipnets. The crew proceeded slowly up each riffle using this technique. Pools were fished by concentrating on fish habitat (e.g., undercut banks, root wads, boulders, etc.) when present or were otherwise fished from side to side in an upstream manner (Barbour et al. 1999). We did not use block nets to enclose each segment of the sample site because this reflects how stream fishes were historically sampled in much of Arkansas. In addition, Simonson and Lyons (1995) found that block nets were not needed to assess species richness and assemblage structure when electrofishing using a single, upstream pass. Paller (1995) also indicated that blocking off habitat units was unnecessary because catch rates in segments with block nets were not greater than catch rates in other segments, indicating that if fish were avoiding the electrofishing gear it was not only in an upstream direction. Furthermore, we were trying to develop a sampling protocol that may be utilized by agency personnel (e.g., rapid bioassessment protocols

(Barbour et al. 1999)) and research biologists. Because of this, we thought that using block nets would reduce the applicability of our findings for most users in this region.

At least one author was present during each sampling episode to ensure sampling consistency. Fish collected from each stream segment were kept separate, preserved in 10% formalin, and identified to species in the laboratory. Age-0 fishes (except lamprey ammocoetes; Family Petromyzontidae) were removed from all segment samples because they could not be accurately identified. Also, potential bias may result if juveniles are included when estimating community attributes (Peterson and Rabeni 1995). Lamprey ammocoetes were not omitted from our samples because lamprey adults in Arkansas migrate to streams in late winter and early spring to spawn (Robison and Buchanan 1988), times of the year when Arkansas' streams are not typically sampled. Therefore, adults are rarely collected during sampling episodes, and including ammocoetes in our samples indicates past adult lamprey use of a stream site.

Sampling Effort Relations with Fish Assemblage Characters

We evaluated the relations between sampling effort and estimates of species richness and assemblage structure. Species richness was calculated as the total number of species collected. The effect of sampling effort on assemblage structure was assessed by determining the community similarity of accumulated segments and the entire sample reach. Community similarity was calculated using the Simplified Morisita's Index (Horn 1966). We used the Simplified Morisita's Index to quantify community similarity because Krebs (1989) found it less biased than other similarity indices that are calculated using proportional data. The index is calculated as

$$C_H = \frac{2\sum_{i=1}^{n}(p_{ij}p_{ik})}{\sum_{i=1}^{n}p^2_{ij} + \sum_{i=1}^{n}p^2_{ik}}$$

where C_H is the Simplified Morisita's Index of overlap (i.e., similarity in community structure) between segment j (i.e., the number of accumulated segments) and all segments combined (k), p_{ij} and p_{ik} are the proportions of species i occurring in segments j and k respectively, and n is the number of species collected in all segments combined. The index ranges from 0 (no similarity in community structure) to 1 (complete similarity), and we express this as a percentage (C_H * 100).

All species richness and community similarity evaluations were made using a non-linear regression technique (Gauss-Newton method; SAS Institute Inc. 1999) and were fitted to the negative exponential function

$$y = A - Be^{-Cx}$$

where A, B, and C are parameter coefficients, e is a logarithmic constant, x is the number of MSWs sampled, and y is either the percent of total species collected or percent community similarity. The coefficient A was set at 100.0 in each model because it is the true asymptote in both non-linear relations.

We evaluated the effect of increased sampling effort on species richness and assemblage structure using three different methods. Once for each individual site (n = 16) in an upstream (from segment one to segment 15) and downstream (from segment 15 to segment one) direction, for all upstream data combined (n = 240) and downstream data combined (n = 240), and for all upstream and downstream data combined (n = 480) (note that sample sizes in this case are based on those used for regression analysis in which a

single segment is considered an observation). Data for each site included zero fish when zero segments were sampled. Therefore, each site contained 16 observations.

Before pooling data among streams, we tested whether stream size or species richness affected the estimated number of MSWs needed to obtain 95% of all species and 95% similarity in individual streams. Effects of these two variables could potentially be masked when pooling data from individual streams. To determine if an increased number of stream widths need sampling in larger streams we conducted a linear regression analysis between mean stream width and the distance needed to collect 95% of all species and reach 95% community similarity. To determine if an increased number of stream widths need sampling in streams with higher species richness we conducted a linear regression analysis between species richness and the distance needed to collect 95% of all species and reach 95% community similarity at each site. We used each site's upstream and downstream data in both analyses (n = 30). No significant relation was found in either analysis. Therefore, we concluded that pooling data from streams of different sizes and species richness would not confound further analyses.

We also assessed whether differences occurred in the sampling distances required to collect 95% of all species and reach 95% similarity in upstream and downstream directions. Because the sampling distances required to sample 95% of all species and 95% similarity differed between upstream and downstream data we pooled upstream and downstream data (i.e., n = 480) to evaluate changes in species richness and community similarity with increasing sampling effort. We tested for the effect of sampling effort for both upstream and downstream directions because the presence and location of microhabitat types in riffles and pools may have varied in a random fashion between

upstream and downstream directions. This would simulate streams of similar physiochemical character, but yield different accumulation rates. Thus, we assumed that pooling upstream and downstream data resulted in the best average estimates.

Estimated regression coefficients from the negative exponential model for pooled data (Table 1.2) were used to determine the expected stream length at which 95% of the species were collected and where 95% community similarity occurs. Ninety-five percent confidence intervals were calculated to determine the variability of species richness and community similarity when using the average length needed to collect 95% of all species and reach 95% community similarity in a stream reach. However, because the 95% confidence intervals do not converge as sampling effort increases for this analysis, as would be expected, the confidence intervals calculated in this way may not reflect what is expected from empirical results. Therefore, we also calculated the average number and percent of species that would be missed when sampling only the distance predicted in which 95% of all species would be collected. To do this, we randomly removed the percent of individuals from each fish collection equal to the percent difference between the average distance needed to collect 95% of all species and 75 MSWs. Based on the individuals removed, we determined the number and percent of species that were omitted from each stream's sample. Confidence intervals for the percent of species removed were calculated using arcsin-transformed data (Zar 1999).

We also determined the number of mean bankfull widths and riffle-pool sequences needed to collect 95% of the species collected and reach 95% community similarity. We did this by calculating the number of mean bankfull widths and riffle-pool

sequences that equaled the predicted number of MSWs needed to collect 95% of the species and to reach 95% community similarity.

Habitat Variability

To determine which stream habitat parameter varied least, as useful characteristic for predictive purposes, we tested for variability among MSW, mean bankfull width, and mean riffle-pool length by using the coefficients of variation (CV) for each habitat parameter measured at each site. We tested for differences among the CVs with a single-factor ANOVA (Zar 1999). We used the Tukey test for multiple comparisons when the ANOVA indicated significant differences. Alpha levels were set at 0.05 for both tests.

Taxa Patchiness

We determined local taxa patchiness by calculating the median, and generating box plots, of the proportion of segments in which each taxon occurred. For example, if a taxon occurred in 5 of 15 total segments at a site, it was present in a 0.333 proportion of all segments. That proportion represented one data point in the box plot for that taxon. This information enabled us to assess which taxa were most likely to be missed by sampling a stream length that will on average include both 95% of the species present and represent 95% community similarity (i.e., the average distance needed). Because we found that, on average, 51 MSWs was the distance at which 95% of the species were collected, if a taxon occurred in two out of 15 sample segments (a proportion of 0.133), it had a probability of occurring in less than one sample segment (5 MSWs / 51 MSWs = 0.098) using the average distance needed (51MSWs / 75MSWs = 0.680; 0.680 x 0.133 =

0.090). Thus, those species with a median occurrence of 0.133 have less than a 10% chance (i.e., 0.133 x 0.680 = 0.090) of occurrence, and a median probability of occurring in less than one sample segment (0.090 < 0.098) using the average distance needed to estimate both species richness and community similarity.

Estimates of MSW

All previous analyses were conducted using our field estimates of MSW. This stratified transect approach (Platts et al. 1983) resulted in habitat units with varying transect numbers. Although we attempted to keep transects evenly spaced, habitats with higher stream-width heterogeneity received higher numbers of transects per habitat length. We assume this resulted in accurate estimates of each habitat's average stream width, but may not have perfectly reflected the true MSW for each stream. To assess the accuracy of our field estimates we calculated a length-weighted MSW for each stream using a bootstrap technique (Sokal and Rohlf 1995). The length-weighted MSW was generated by randomly resampling, with replacement, each habitat unit's stream-width measurements so that each habitat unit contained a number of stream width measurements equal to the percent of that habitat's contribution to the sampled stream reach's length. All percentages were rounded to the nearest integer and all habitat units contributed at least one measurement, thus, not all stream sites had exactly 100 stream width measurements. We used a paired t-test to test for differences ($\alpha = 0.05$) between our field and length-weighted MSW estimates for each stream site.

We also determined the number of MSWs needed to collect 95% of all species and reach 95% community similarity using the length-weighted means. To calculate this,

the mean proportional difference between the field and length-weighted MSW estimates was multiplied by the number of MSWs needed to collect 95% of all species and reach 95% community similarity.

In order to aid managers and researchers in determining the sample size needed to estimate the MSW in Ozark Highland streams, we estimated sample sizes using our field stream width measurements and length-weighted stream-width data set. We applied a sample size formula similar to Zar (1999) and Churchill (1999) to both our field and length-weighted stream-width measurements. The sample size formula is

$$n = \frac{(t_{\alpha(2),(v)})^2 (CV)^2}{p^2}$$

where n is the sample size, t is the critical value of the t distribution, α was set at 0.05, v is the degrees of freedom for each sampling unit, CV is the coefficient of variation, and p is the allowable error between the estimated mean and true population mean (unknown). p was arbitrarily set at 5, 10, and 20%. We also determined the percent error between our field and length-weighted MSW estimates and the true population MSW using the mean sample size and coefficient of variation from each MSW estimate. The percent error was determined by rearranging the sample size formula to solve for p.

Results

Sample Sites

The mean distance sampled for all streams was 501.1 m (SD = 164.1), and the range was 246 m (Upshaw Creek) to 790 m (Long Creek). Discharges averaged 0.1464 (SD = 0.1338) cubic meters per second (m^3/s) and ranged from 0.009 m^3/s at Upshaw

Creek to 0.387 m^3/s at Mud Creek (Table 1.3). We collected 50 total species among all sample sites; the most species occurred in Long Creek (26); the fewest occurred in Dry Creek (9).

Sampling Effort Relations with Fish Assemblage Characters

We found no significant relation between MSW and the number of MSWs needed to collect 95% of all species (p-value = 0.628; $r^2 = 0.009$) and obtain 95% community similarity (p-value = 0.870; $r^2 = 0.001$). We also found no significant effect of species richness on the sampling effort needed to obtain 95% of all species (p-value = 0.519; $r^2 = 0.015$) or 95% community similarity (p-value = 0.201; $r^2 = 0.058$). Nonlinear regression results from upstream and downstream data sets indicated that the average distance needed to collect 95% of all species (upstream = 40.8 MSWs; downstream = 64.2 MSWs) and reach 95% community similarity (upstream = 16.2 MSWs; downstream = 10.9 MSWs) differed between data sets. Therefore, we pooled data from all sites for our final non-linear regression analysis.

With data from all sites, in both upstream and downstream directions incorporated, the model predicted that 95% of the species will be sampled at a stream length of 51.0 MSWs (Figure 1.2). The upper and lower 95% confidence intervals of the percent of species accumulated at 51.0 MSWs generated by the nonlinear regression were 100 and 72.5% respectively. Based on the proportional random removal of species from the actual data sets, the average number of species missed at 51.0 MSWs was 1.0 (n = 15; SD = 1.1). This represents a percent of 5.6 with upper and lower confidence intervals of 7.0 and 4.2% respectively. The estimated average number of MSWs needed to collect

95% of the species equaled 16.1 mean bankfull widths or 5.7 riffle-pool sequences. The model also predicted that, on average, 13.4 MSWs were needed to reach 95% similarity between the predicted and sampled fish communities (Figure 1.3). By sampling 13.4 MSWs one can be 95% confident they will reach between 100 and 74.8% community similarity. An estimated 4.2 mean bankfull widths or 1.5 riffle-pool sequences need sampling in order to reach 95% community similarity. Table 1.2 lists the estimated coefficients for the final non-linear regression analysis.

Habitat Variability

The habitat variability analysis indicated that among sample streams the mean CVs were 43.2 (SD = 12.1) for MSW, 34.3 (SD = 21.3) for mean bankfull width, and 95.5 (SD = 24.6) for mean riffle-pool length (Table 1.3). Significant differences existed among habitat parameter CVs (p-value < 0.01), and multiple comparisons revealed significant differences between CVs of mean bankfull width and mean riffle-pool length, and CVs of MSW and mean riffle-pool length. No significant difference was detected between MSW and mean bankfull width CVs.

Taxa Patchiness

Zero families, six genera and 12 species showed a median probability of occurring in less than one segment when sampling 51.0 MSWs (Figure 1.4) (Table 1.4). Genera with less than a 10% chance of occurring in at least one segment in a stream reach were *Ameiurus*, *Cyprinus*, *Labidesthes*, *Micropterus*, *Moxostoma*, and *Percina*. Species with a probability of occurring in less than one segment using the average

estimated length were *Cyprinus carpio*, *Pimephales promelas*, *Moxostoma carinatum*, *Moxostoma duquesnei*, *Moxostoma erythrurum*, *Ameiurus melas*, *Fundulus catenatus*, *Labidesthes sicculus*, *Lepomis microlophus*, *Micropterus punctulatus*, *Micropterus salmoides*, and *Percina caprodes*. Most taxa occurred in only one or two stream segments in at least one site when present at a site (Figure 1.4). Therefore, most taxa have the potential to be missed when sampling a length of 51.0 MSWs. However, one cannot predict *a priori* which taxa will be missed when sampling a site.

Estimates of MSW

The mean field MSW estimate was 6.68 m (SD = 2.26) and the mean length-weighted MSW estimate was 7.40 m (SD = 2.50). The length-weighted MSW estimate was greater than the field MSW estimate for all stream sites except one. This resulted in a significant difference ($p < 0.01$) between the field and length-weighted MSW estimates for our sample sites.

Our field MSW estimates were on average 0.903 proportion of the length-weighted MSW estimates. Therefore, 46.1 length-weighted MSWs (51.0 * 0.903) are needed to collect 95% of all species and 12.1 length-weighted MSWs (13.4 * 0.903) are needed to obtain 95% community similarity.

Using the field stream-width measurements, the mean sample sizes needed to obtain a MSW estimate within 5, 10, and 20% of the true population MSW, with 95% confidence, are 326, 82, and 21 respectively. Using the length-weighted stream-width measurements the mean sample sizes needed to obtain a MSW estimate within 5, 10, and 20% of the true population MSW, with 95% confidence, are 279, 70, and 18 respectively

(Table 1.5). At a 95% confidence level our field MSW estimates were, on average, within 13.3% of the true population mean, while the length-weighted MSW estimates were within 8.0%.

Discussion

Sampling 46.1 – 51.0 MSW in wadeable Ozark Highland streams using backpack electroshocking will, on average, collect at least 95% of all species and nearly perfectly reflect fish-assemblage structure. Species richness and assemblage structure each represent important fish-assemblage attributes when studying fish assemblages and determining stream health by using biological indicators (Karr et al. 1986). When sampling the average distance needed, the species most likely to be missed cannot be determined *a priori*. Also, we recommend scaling sampling distances by MSW because of its relatively low variability compared to riffle-pool sequence length and its application ease when compared to mean bankfull width (Simonson et al. 1993).

The average sampling distance needed to characterize stream fish assemblages in wadeable, Ozark Highland streams was close to other estimated sampling distances needed to characterize key stream-fish-assemblage attributes (Lyons 1992; Paller 1995; Angermeier and Smogor 1995; Barbour et al. 1999; Patton et al. 2000). Community similarity in Arkansas' Ozark Highlands, on average, accumulated in slightly less distances (12.0 - 13.4 MSW) than the 15 - 20 MSWs reported by Angermeier and Smogor (1995). Although Angermeier and Smogor (1995) used a different similarity index than us, results should be generally comparable (Krebs 1989). Our estimated 46.1 - 51.0 MSWs needed to collect 95% of the total species is greater than the 35 MSWs

recommended by Lyons (1992) and the 40 MSWs recommended for the U.S. Environmental Protection Agency's (USEPA) Environmental Monitoring and Assessment Program (EMAP) (Klemm and Lazorchak 1995). The 46.1 – 51.0 MSWs needed in our streams to capture 95% of all species encompassed the maximum number of MSWs (50) that need sampling by electrofishing to collect 100% of all species in Wyoming's Great Plains streams (Patton et al. 2000). Also, our average distance needed to collect 95% of all species at eight of our sample sites was greater than the maximum sampling distance of 300 m Meador et al. (1993) recommends for sampling wadeable streams for the NAWQA program. Conversely, the average sampling distance is within the 45 to 90 MSWs reported by Angermeier and Smogor (1995). Angermeier and Smogor (1995) indicated that the particular stream that needed 90 MSWs sampled contained little habitat heterogeneity. They suggested that habitat homogeneity influenced population density and species spatial discontinuity (i.e., patchiness) which resulted in a greater sampling effort needed to characterize the stream-fish assemblage in their sample site with homogeneous habitat. Because mountainous terrain and steep gradients characterize Arkansas' Ozark Highlands (Robison and Buchanan 1988), streams within this region have observable habitat heterogeneity (i.e., riffle-pool sequences), even in disturbed systems. This may explain why 46.1 – 51.0 MSWs falls in the lower portion of Angermeier and Smogor's (1995) range of estimated lengths.

Our results suggest there are mechanisms causing some species' distributions to be patchy. Bart (1989) studied habitat use of stream fishes in an Ozark stream and found that most fish species used a variety of habitats and were not specific to riffle or pool habitats. Gorman (1988) also found habitat use for certain Ozark minnow species to be

variable. Although these results suggest that most Ozark fish species in streams may not specifically use riffle or pool habitats, further information is required to determine if mechanisms causing discontinuous species distributions in Ozark streams are similar to those found in a Virginia stream (Angermeier and Smogor 1995).

Based on the nonlinear regression model, when applying the average sampling distance needed one can be 95% confident that they have collected at least 72.5% of the species and have reached at least 74.8% community similarity in a stream reach. Although some variability was expected, these confidence intervals from the regression model seem unlikely to reflect reality because all regression confidence intervals widen with distance from the mean coordinates (Sokal and Rohlf 1995). Consequently, the confidence intervals in our model never converge, as expected, as the number of stream segments reaches its maximum. In addition, the average percent of species missed was only 5.6 with an upper confidence interval of 7.0% when simulating a sample of 46.1 – 51.0 MSWs using a random removal procedure. Missing 7.0% is equivalent to a lower bound of 93%, which is much higher, and appears more realistic based on distributions of our data, than what was predicted by the regression model's lower confidence interval. Nonetheless, there was variability in the rates at which species richness and community similarity accumulated.

Our goal to determine a single sampling distance for application across the Ozark Highlands ecoregion likely contributed to some of the variability identified in species richness and community similarity accumulations. We sampled streams with observable variation in disturbance levels, dominant substrate type (i.e., bedrock and alluvial), degree of spring influence, and gradient, among other things. Though these variables

were not quantified and their degree of influence is unknown, we believe they contributed to some of the variability found in our model. Due to the observed variability, when very precise estimates of species richness are needed our findings may not be useful for wadeable, Ozark Highland streams. We tried to limit this variability by utilizing existing sampling protocols for the region, using a single gear type, sampling only wadeable streams, and using a standardized sample length based on stream size as recommended by Lyons (1992). We also sampled during summer months when streams are at or near base flow, a time when most agencies and other researchers conduct their sampling for community assessment purposes. Despite these efforts, some variability in number of MSWs needed to estimate species richness and community similarity existed for our sample sites.

Estimates of MSW may have also contributed to the variability in our model. Although we assumed our field estimates of MSW were adequate, when compared to a length-weighted MSW they differed significantly. Also, our field MSW estimates were only within 13.3% of the true population MSW. This means that our MSW estimate constitutes a 26.6% range around the true population mean, which is greater than the 9.7% difference found between our field and length-weighted MSW estimates. Inaccurate MSW estimates can add additional variability in models using MSW as an independent variable. Therefore, increased accuracy in MSW estimates should minimize a potential source of variability. The level of accuracy will depend on study objectives, which has rarely been addressed in studies relying on MSW estimates. Simonson et al. (1994) addressed the accuracy of MSW estimates and found approximately 20 MSW measurements spaced 2 MSWs apart sufficient to generate MSW estimates within 5% of

the true MSW. Although, their transect spacing was dependent on only 15 – 20 preliminary stream-width measurements. We found Ozark Highland streams to contain more stream-width variability than Wisconsin streams (Simonson et al. 1994), therefore, requiring increased sample sizes to obtain accurate MSW estimates. Also, when using MSW to standardize sampling effort it should be acknowledged that unless numerous samples of wetted width are taken, the potential errors in MSW estimates will remain relatively high in streams with variable stream width. Despite this, we recommend using MSW for determining sampling distances because even though MSW and bankfull width variability were not significantly different, MSW scales sampling distance to stream size and is more discernable and easier to measure than bankfull width (Simonson et al. 1993; Wang et al. 1996). Also, at least one government agency has ceased routinely measuring bankfull width because of the difficulty in determining the precise location of where bankfull measurements should be taken (Simonson et al. 1993; Simonson et al. 1994).

The effect of variability in accumulation rates of fish-assemblage characteristics could possibly be minimized using a variety of approaches. As mentioned above, accurate MSW estimates may minimize the variability. Many methods of transect spacing (Brown and Austen 1996) and transect clustering (Platts et al. 1983) are available. Simonson et al. (1994) recommends systematic transect spacing according to MSW because of time limitations and periodicity of stream morphological characters. Appropriate stratification of stream types may also reduce the variability. For example, stratifying by drainage basin, level of degradation, elevation, etc. may account for much of the variability. For the purposes of comparing sites to reference conditions the best approach may be to examine the effect of sampling effort on fish-assemblage characters

in least-disturbed streams only. This may reduce the variability associated with site condition and could potentially result in a single, regional sampling distance. This would allow comparisons between impacted and unimpacted fish assemblages, and may be the most appropriate approach. It must be remembered that sampling effort and fish assemblage comparisons are almost always dictated by study objectives. For example, some studies concerning biotic indices have shown that increased sampling reduces variability in index scores (Angermeier and Karr 1986; Paller et al. 1996). It is suggested that standard sampling protocols that reach the species-richness asymptote be used because slight variation around that distance will have a minimal effect on the number of species collected (Paller et al. 1996). Also, sampling less than the asymptotic distance may result in failure of taxonomic metrics (e.g., number of sunfish species) to reflect stream health if not all species are collected. Local species presence and abundance at our sample sites may have also attributed to the variability in our model. These assemblage characters in Ozark Highland streams may be determined by factors influencing habitat heterogeneity. Species presence and abundance in other regions have been shown to change seasonally (Taylor et al. 1996; Tripe and Guy 1999; Peterson and Rabeni 2001) and annually (Schlosser 1985; Poff and Allan 1995) as flow regimes change. These same mechanisms may affect availability and abundance of habitat and refugia in Ozark Highland streams, which may cause shifts in fish assemblage structure and abundance (Dewey 1981). Although, in Illinois, Schlosser (1985) reported that only juvenile abundance changed, while adult abundance remained stable between years having different flow regimes. Matthews (1986) also suggests Ozark fish assemblages

are stable across seasons. We attempted to minimize any potential seasonal effects by conducting summer sampling at conditions that reflected base stream flow.

Taxa showing a low probability of occurrence at our sample sites may also have affected the rate at which species richness accumulated depending on the sample segment in which they were collected. Various factors may contribute to species occurrence and patchiness in wadeable, Ozark Highland streams (see Jackson et al. (2001)). Some species of low abundance collected at our sites are generally found in larger streams than the streams we sampled. For example, *Moxostoma* species prefer medium-sized streams and migrate to higher-gradient streams to spawn (Robison and Buchanan 1988). The few individuals we encountered may have been remnants from the spawning season, or were collected in stream segments that provided rare, suitable habitat for those individuals. Schooling species may also show a low probability of occurrence since all individuals may be collected together in one stream segment. Another factor influencing species accumulation rates may be that the Ozark Highlands ecoregion in Arkansas comprises three major drainages (i.e., the Illinois River, White River, and Black River), and individual species may occur in only one drainage (e.g., *Luxilus zonatus* is confined to the Black River drainage). Though mechanisms driving species abundance are undetermined for Ozark Highland streams (see Matthews (1982)), we think these mechanisms may also contribute to the variability we found in the average distance needed to characterize stream-fish assemblages in the region. Also, because differences exist in sampling distances needed to estimate fish-assemblage characters, mechanisms determining local abundance and patchiness may differ between regions in the U.S. (Poff 1997; Strange 1999).

Although results from this study provide useful information concerning stream-fish-assemblage characterization, persons using these findings should proceed with knowledge of critical assumptions and certain limitations. Because of interregional differences in species-area relations, applying these results outside of Arkansas' Ozark Highlands is inappropriate in the absence of testing for similar species accumulation rates. In addition, all sampling gears are biased, and our samples reflect biases associated with backpack electrofishing. However, backpack electrofishing is a well-suited and commonly used technique for sampling wadeable streams (Barbour et al. 1999). Additionally, Wiley and Tsai (1983) found a particular electrofishing gear was more consistent than seines, another popular gear type for sampling wadeable streams, when estimating fish populations. Patton et al. (2000) also found shorter sampling distances were required to collect all species using electrofishing gear as opposed to seining in Great Plains streams.

Although we found that scaling the size of the stream using MSW as a measure of stream size worked well in this study, this relation may not hold for all streams. Paller (1995) and Patton et al. (2000) found that shorter distances were needed in larger streams to obtain maximum and 90% species richness respectively. Lyons (1992) also suggested an absolute sampling distance may be appropriate in Wisconsin streams. Angermeier and Smogor (1995) suggested that an adequate sampling distance is, to some degree, related to habitat homogeneity and species discontinuity. We found no relation between stream size and the number of MSWs needed to estimate species richness. This may have resulted from sampling a limited range of stream sizes that generally had well-distributed riffle-pool sequences, which is characteristic of the region. We only sampled wadeable

streams and species-effort relations may change in Arkansas' non-wadeable streams due to increased species richness (Ebert et al. 1992; Matthews and Robison 1998). Also, alternative gears are required to sample non-wadeable streams and rivers. These factors may yield different rates of species accumulation and community similarity (Wiley and Tsai 1983), warranting further research. Therefore, our results should be useful for people sampling with backpack electroshockers in wadeable, Ozark Highland streams.

In summary, we found that sampling 46.1 – 51.0 MSWs in wadeable Ozark Highlands streams in Arkansas will, on average, collect at least 95% of the fish species present and adequately characterize the relative proportion of species in the fish assemblage. Although some variability in the completeness of the sample may result from sampling a reach of 46.1 – 51.0 MSWs, such a protocol will improve the chances to consistently characterize fish assemblages for the purpose of comparing stream-fish assemblages in the ecoregion. Our results should provide regional professionals with valuable information when setting their study objectives and forming their sampling protocols relative to available resources. Though individual study or management objectives may not require an accurate assessment of species or taxonomic richness, knowledge of what information a given sampling episode may potentially yield is important and should be recognized. Further, our results should ultimately contribute to the body of literature that will answer the geographically broader question regarding how much sampling is enough to characterize stream-fish assemblages.

Acknowledgements

This manuscript was funded through a sub-contract agreement with Tetra Tech, Inc. via the U.S. Environmental Protection Agency contract # 68-C98-111 Work Assignment 1-10. We thank C. Okiror, J. Poole, T. Hungerford, L. Lackey, and S. Mutagyera for help with data collection and data entry. We also thank N. Stone, E. Buckner, S. Lochmann, W. Keith, P. Crocker, R. Smogor, and M. Barbour for critical reviews and comments regarding earlier drafts of this chapter.

References

ADPCE. 1996. Arkansas water quality inventory report. Arkansas Department of Pollution Control & Ecology, Little Rock, Arkansas.

Angermeier, P. L. and J. R. Karr. 1986. Applying an index of biotic integrity based on stream-fish communities: considerations in sampling and interpretation. North American Journal of Fisheries Management 6:418-429.

Angermeier, P. L. and I. J. Schlosser. 1989. Species-area relationships for stream fishes. Ecology 70:1450-1462.

Angermeier, P. L. and R. A. Smogor. 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. Canadian Journal of Fisheries and Aquatic Sciences 52:936-949.

Arrhenius, O. 1921. Species and area. Journal of Ecology 9:90-91.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in wadeable streams and rivers: periphyton, benthic macroinvertebrates, and fish, Second edition. U.S. Environmental Protection Agency; Office of Water; Washington D.C. EPA 841-B-99-002.

Bart, H. L. Jr. 1989. Fish habitat association in an Ozark stream. Environmental Biology of Fishes 24:173-186.

Brown, M. L. and D. J. Austen. 1996. Data management and statistical techniques. Pages 17-62 *in* B. R. Murphy and D. W. Willis, editors. Fisheries Techniques. American Fisheries Society, Bethesda, Maryland.

Churchill, G. A. Jr. 1999. Marketing research: methodological foundations, 7th ed. The Dryden Press, Orlando, Florida.

Dewey, M. R. 1981. Seasonal abundance, movement and diversity of fishes in an Ozark stream. Proceedings Arkansas Academy of Science 35:33-39.

Ebert, D. J., S. P. Filipek, and K. M. Russell. 1992. Stream habitat analysis and instream flow assessment: a state-federal effort in Arkansas. U.S.Fish & Wildlife Service Biological Report 90:43-44.

Gallagher, A. S. and N. J. Stevenson. 1999. Streamflow. Pages 149-157 *in* M. B. Bain and N. J. Stevenson, editors. Aquatic habitat assessment: common methods. American Fisheries Society, Bethesda, Maryland.

Giese, J., B. Keith, M. Maner, R. McDaniel, and B. Singleton. 1987. Physical, chemical, and biological characteristics of least-disturbed reference streams in Arkansas' ecoregions: data analysis. Arkansas Department of Pollution Control & Ecology, Little Rock.

Gleason, H. A. 1922. On the relation between species and area. Ecology 3:158-162.

Gorman, O. T. 1988. The dynamics of habitat use in a guild of Ozark minnows. Ecological Monographs 58:1-18.

Horn, H. S. 1966. Measurement of "overlap" in comparative ecological studies. American Naturalist 100:419-424.

Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities - the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences 58:157-170.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running water: a method and its rationale. Illinois Natural History Survey, Champaign, Illinois. Special Publication 5.

Klemm, D. J. and J. M. Lazorchak. 1995. Environmental monitoring and assessment program - surface waters: field operations and methods for measuring the ecological conditions of wadeable streams. Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH. EPA/620/R-94/004.

Krebs, C. J. 1989. Ecological methodology. Harper & Row, Publishers, New York, New York.

Lyons, J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. North American Journal of Fisheries Management 12:198-203.

Matthews, W. J. 1982. Small fish community structure in Ozark streams: structured assembly patterns or random abundance of species. The American Midland Naturalist 107:42-54.

Matthews, W. J. 1986. Fish faunal structure in an Ozark stream: stability, persistence, and a catastrophic flood. Copeia 1986:388-397.

Matthews, W. J. and H. W. Robison. 1998. Influence of drainage connectivity, drainage area and regional species richness on fishes of the interior highlands in Arkansas. The American Midland Naturalist 139:1-19.

Meador, M. R., T. F. Cuffney, and M. E. Gurtz. 1993. Methods for sampling fish communities as part of the National Water Quality Assessment. U.S. Geological Survey Open File Report, Raleigh, North Carolina. USGS/OPR 93-104.

Paller, M. H. 1995. Relationships among number of fish species sampled, reach length surveyed, and sampling effort in South Carolina Coastal Plain streams. North American Journal of Fisheries Management 15:110-120.

Paller, M. H., M. J. M. Reichert, and J. M. Dean. 1996. Use of fish communities to assess environmental impacts in South Carolina coastal plain streams. Transactions of the American Fisheries Society 125:633-644.

Patton, T. M., W. A. Hubert, F. J. Rahel, and K. G. Gerow. 2000. Effort needed to estimate species richness in small streams in the Great Plains in Wyoming. North American Journal of Fisheries Management 20:394-398.

Peterson, J. T. and C. F. Rabeni. 1995. Optimizing sampling effort for sampling warmwater stream fish communities. North American Journal of Fisheries Management 15:528-541.

Peterson, J. T. and C. F. Rabeni. 2001. The relation of fish assemblages to channel units in an Ozark stream. Transactions of the American Fisheries Society 130:911-926.

Plafkin, J. L., M. T. Barbour, S. K. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C. EPA 440-4-89-001.

Platts, W. S., W. F. Megahan, and G. W. Minshall. 1983. Methods for evaluating stream, riparian, and biotic conditions. U.S. Forest Service, Intermountain Forest and Range Experiment Station, Ogden, Utah.

Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391-409.

Poff, N. L. and J. D. Allan. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. Ecology 76:606-627.

Robison, H. W. and T. M. Buchanan. 1988. Fishes of Arkansas. The University of Arkansas Press, Fayetteville, Arkansas.

Roper, B. B. and D. L. Scarnecchia. 1995. Observer variability in classifying habitat types in stream surveys. North American Journal of Fisheries Management 15:49-53.

SAS Institute Inc. 1999. SAS OnlineDoc, Version 8. SAS Institute Inc., Cary, North Carolina.

Schlosser, I. J. 1985. Flow regime, juvenile abundance, and the assemblage structure of stream fishes. Ecology 66:1484-1490.

Simonson, T. D. and J. Lyons. 1995. Comparison of catch per effort and removal procedures for sampling stream fish assemblages. North American Journal of Fisheries Management 15:419-427.

Simonson, T. D., J. Lyons, and P. D. Kanehl. 1993. Guidelines for evaluating fish habitat in Wisconsin streams. U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station, St. Paul, MN. General Technical Report NC-164.

Simonson, T. D., J. Lyons, and P. D. Kanehl. 1994. Quantifying fish habitat in streams: transect spacing, sample size, and a proposed framework. North American Journal of Fisheries Management 14:607-615.

Smith, K. G., R. S. Dzur, D. G. Catanzaro, M. E. Garner, and W. F. Limp. 1998. The Arkansas GAP analysis project: final report. Center for Advanced Spatial Technologies (CAST), Fayetteville, Arkansas.

Sokal, R. R. and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research, 3rd ed. W. H. Freeman and Company, New York.

Strange, R. M. 1999. Historical biogeography, ecology, and fish distributions: conceptual issues for establishing IBI criteria. Pages 65-78 *in* T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Inc., Boca Raton, Florida.

Taylor, C. M., M. R. Winston, and W. J. Matthews. 1996. Temporal variation in tributary and mainstem fish assemblages in a Great Plains stream system. Copeia 1996:280-289.

Tripe, J. A. and C. S. Guy. 1999. Spatial and temporal variation in habitat and fish community characteristics in a Kansas Flint Hills stream. Ecology of Freshwater Fish 8:216-226.

Wang, L., T. D. Simonson, and J. Lyons. 1996. Accuracy and precision of selected stream habitat estimates. North American Journal of Fisheries Management 16:340-347.

Weih, R. 2001. Natural state digital database: version 1.0 Beta 5, UTM Zone 15, NAD27. Arkansas Forest Resource Center, University of Arkansas-Monticello, Monticello, Arkansas.

Wiley, M. L. and C. Tsai. 1983. The relative efficiencies of electrofishing vs. seining in Piedmont streams of Maryland. North American Journal of Fisheries Management 3:243-253.

Zar, J. H. 1999. Biostatistical analysis, 4th ed. Prentice Hall, Upper Saddle River, New Jersey.

Table 1.1. Watershed characteristics of the 15 study sites in Arkansas' Ozark Highlands.

	Location	Watershed Size	Land Use		
Stream	County	km^2	% Forested	% Agriculture	% Urban
Big Creek	Fulton	16.47	51.79	48.18	0.00
Brush Creek	Washington	57.53	20.75	79.24	0.00
Clear Creek	Washington	28.84	0.00	93.55	3.80
Diles Creek	Randolph	25.02	88.84	11.16	0.00
Dry Creek	Carroll	23.07	46.95	52.57	0.47
Greasy Creek	Marion	31.57	59.19	40.80	0.00
Hampton Creek	Marion	70.64	43.62	56.48	0.00
Harding Creek	Lawrence	11.76	79.73	20.67	0.00
Long Creek	Carroll	73.56	79.50	20.54	0.00
Mill Creek	Stone	16.37	71.25	28.75	0.00
Mud Creek	Washington	28.31	9.73	76.66	13.62
North Big Creek	Sharp	49.36	33.35	66.64	0.00
North Sylamore Creek	Stone	118.18	99.18	0.83	0.00
Tuttle Branch	Washington	13.58	42.52	57.52	0.00
Upshaw Creek	Randolph	9.11	74.73	25.20	0.00

Table 1.2. Parameter coefficients for the species richness and community similarity models. Both models are negative exponential models of the form $y = A - Be^{-Cx}$.

Coefficient	Species Richness	Community Similarity
A	100.0000	100.0000
B	90.5967	98.5001
C	0.0568	0.2220

Table 1.3. Habitat parameters (means and CVs) and discharge (m^3/s) measured at the 15 study sites in Arkansas' Ozark Highlands. All variables were measured immediately prior to fish sampling.

	Wetted Width (m)		Bankfull Width (m)		Habitat Length (m)		Discharge
Stream	Mean	CV	Mean	CV	Mean	CV	m^3/s
Big Creek	3.80	45	17.71	36	17.82	111	0.004
Brush Creek	8.68	51	18.21	25	46.00	89	0.169
Clear Creek	7.84	24	16.46	44	31.84	97	0.245
Diles Creek	9.08	34	22.79	32	32.27	80	0.140
Dry Creek	6.51	29	12.71	14	31.20	61	0.367
Greasy Creek	4.77	30	13.47	101	22.37	68	0.023
Hampton Creek	8.14	46	63.75	36	38.22	83	0.263
Harding Creek	4.99	34	11.61	28	18.02	122	0.070
Long Creek	10.53	41	15.80	21	48.88	121	0.296
Mill Creek	4.89	55	15.43	25	21.83	72	0.066
Mud Creek	9.98	37	32.86	50	53.96	97	0.387
North Big Creek	5.68	67	36.15	11	28.96	152	0.040
North Sylamore Creek	6.86	53	26.86	21	29.46	71	0.104
Tuttle Branch	5.21	44	13.44	36	30.57	104	0.011
Upshaw Creek	3.28	58	10.79	34	14.03	104	0.009
Mean*		43.3z		34.3z		95.5y	0.1464
Standard Deviation		12.1		21.3		24.6	0.1338

* Identical letters indicate no significant difference ($\alpha = 0.05$) between variables

Table 1.4. Number of times each taxon was collected at a study site and median probability of occurrence for all taxa collected. The last column indicates which taxa have less than a 10% chance of occurrence when sampling 51 MSW.

Taxon	n	Median	<10%
Petromyzontidae	2	0.23	
Ammocoete	2	0.23	
Ammocoete	2	0.23	
Cyprinidae	84	0.33	
Campostoma	15	0.80	
Campostoma anomalum	15	0.80	
Cyprinus	2	0.13	x
Cyprinus carpio	2	0.13	x
Luxilus	15	0.33	
Luxilus cardinalis	1	0.27	
Luxilus chrysocephalus	3	0.20	
Luxilus pilsbryi	7	0.53	
Luxilus zonatus	4	0.27	
Nocomis	8	0.37	
Nocomis bigutattus	8	0.37	
Notropis	17	0.27	
Notropis boops	8	0.17	
Notropis nubilus	7	0.47	
Notropis telescopus	2	0.20	
Phoxinus	5	0.40	
Phoxinus erythrogaster	5	0.40	
Pimephales	11	0.33	
Pimephales notatus	9	0.33	
Pimephales promelas	2	0.07	x
Semotilus	11	0.20	
Semotilus atromaculatus	11	0.20	
Catostomidae	18	0.20	
Catostomus	1	0.60	
Catostomus commersoni	1	0.60	
Erimyzon	5	0.40	
Erimyzon oblongus	5	0.40	
Hypentelium	9	0.20	
Hypentelium nigricans	9	0.20	
Moxostoma	3	0.07	x
Moxostoma carinatum	1	0.07	x
Moxostoma duquesnei	1	0.07	x
Moxostoma erythrurum	1	0.07	x
Ictaluridae	27	0.27	
Ameiurus	15	0.13	x
Ameiurus melas	4	0.07	x
Ameiurus natalis	11	0.27	
Noturus	12	0.47	

Table 1.4. Continued.

Taxon	n	Median	<10%
Noturus albater	1	0.20	
Noturus exilis	11	0.47	
Aphredoderidae	2	0.27	
Aphredoderus	2	0.27	
Aphredoderus sayanus	2	0.27	
Fundulidae	20	0.23	
Fundulus	20	0.23	
Fundulus catenatus	7	0.13	x
Fundulus olivaceus	13	0.33	
Poeciliidae	4	0.27	
Gambusia	4	0.27	
Gambusia affinis	4	0.27	
Atherinidae	1	0.07	x
Labidesthes	1	0.07	x
Labidesthes sicculus	1	0.07	x
Cottidae	11	0.53	
Cottus	11	0.53	
Cottus carolinae	9	0.53	
Cottus hypselurus	2	0.57	
Centrarchidae	58	0.27	
Ambloplites	7	0.33	
Ambloplites ariommus	3	0.20	
Ambloplites constellatus	4	0.37	
Lepomis	34	0.60	
Lepomis cyanellus	13	0.60	
Lepomis gulosus	2	0.17	
Lepomis macrochirus	5	0.27	
Lepomis megalotis	12	0.67	
Lepomis microlophus	1	0.07	x
Lepomis punctatus	1	0.27	
Micropterus	17	0.13	x
Micropterus dolomieu	6	0.23	
Micropterus punctulatus	5	0.07	x
Micropterus salmoides	6	0.13	x
Percidae	48	0.53	
Etheostoma	45	0.60	
Etheostoma blennioides	6	0.43	
Etheostoma caeruleum	13	0.80	
Etheostoma flabellare	6	0.60	
Etheostoma punctulatum	4	0.33	
Etheostoma spectabile	15	0.60	
Etheostoma zonale	1	0.27	
Percina	3	0.07	x

Table 1.4. Continued.

Taxon	n	Median	<10%
Percina caprodes	3	0.07	x

Table 1.5. Sample sizes needed to obtain MSW estimates within 5, 10, and 20% of the true population MSW, with 95% confidence, using field and length-weighted stream-width measurements.

	Field Estimate					Length-Weighted Estimate				
Stream	n	CV	5%	10%	20%	n	CV	5%	10%	20%
Big Creek	42	45	336	84	21	99	41	265	66	17
Brush Creek	33	51	429	107	27	99	48	363	91	23
Clear Creek	51	24	90	22	6	99	22	76	19	5
Diles Creek	52	34	191	48	12	99	33	171	43	11
Dry Creek	43	29	140	35	9	102	27	115	29	7
Greasy Creek	16	30	164	41	10	99	31	151	38	9
Hampton Creek	42	46	347	87	22	100	43	291	73	18
Harding Creek	54	34	182	46	11	100	36	204	51	13
Long Creek	49	41	269	67	17	100	37	216	54	13
Mill Creek	42	55	485	121	30	100	54	459	115	29
Mud Creek	36	37	225	56	14	100	28	123	31	8
North Big Creek	49	67	717	179	45	101	64	645	161	40
North Sylamore Creek	46	53	451	113	28	101	47	348	87	22
Tuttle Branch	40	44	317	79	20	102	40	252	63	16
Upshaw Creek	43	58	539	135	34	101	56	494	123	31
Mean	42.5	43.2	325.5	81.4	20.3	100.1	40.5	278.2	69.5	17.4
SD	9.4	12.1	173.1	43.3	10.8	1.1	11.8	159.3	39.8	10.0

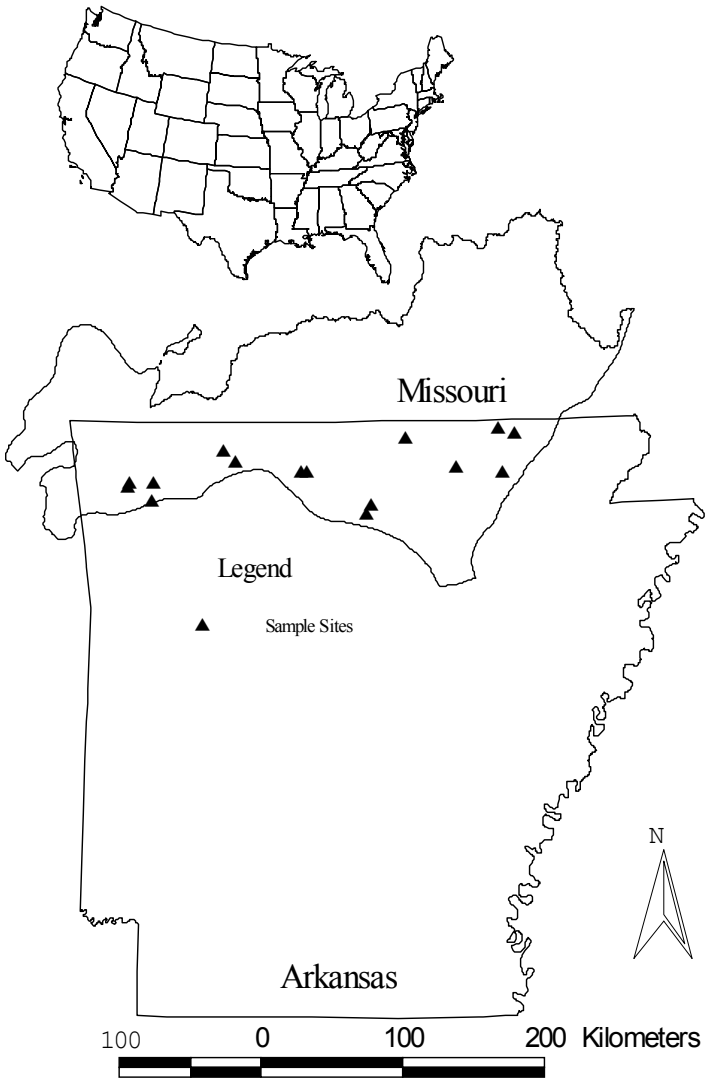


Figure 1.1. The location of our 15 study sites within Arkansas' Ozark Highlands ecoregion.

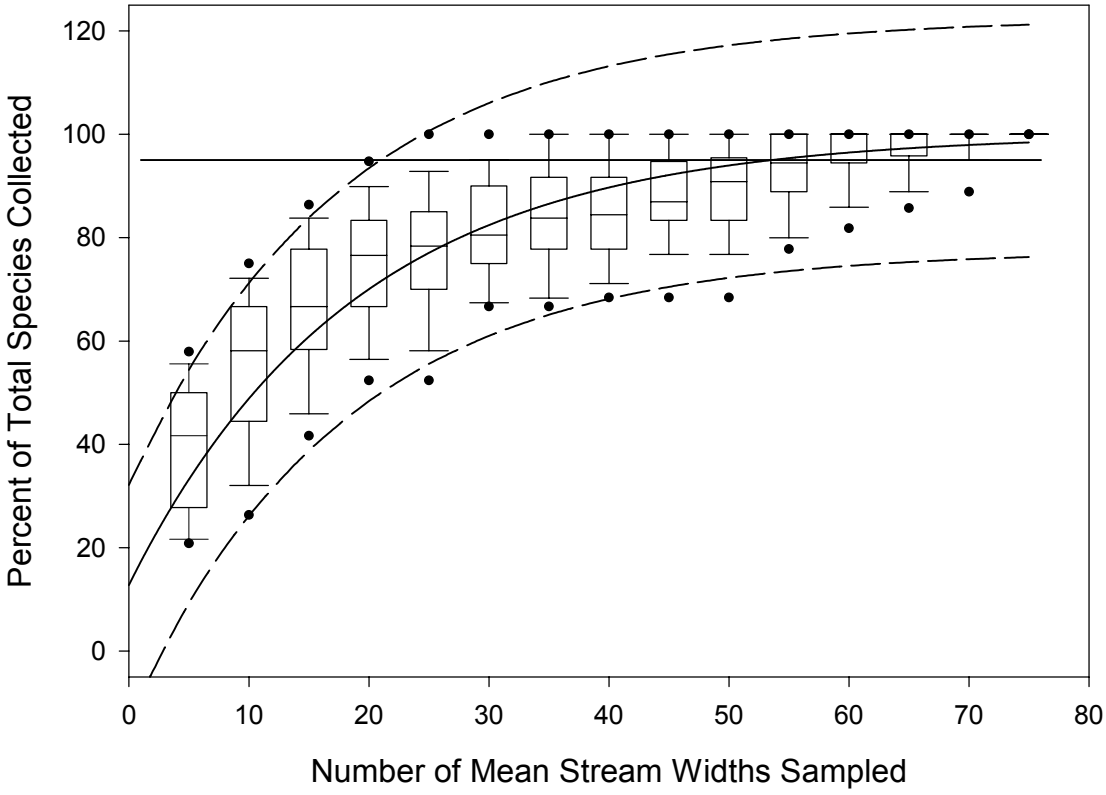


Figure 1.2. Box plots and best-fit line representing the percentage of total species collected every five mean stream widths. Middle line in box plots represents the median, box edges are the 25th and 75th percentiles, t-bars are the 10th and 90th percentiles, and dots represent the 5th and 95th percentiles. Predicted line $y = 100.0000\text{-}90.5967e^{-0.0568x}$ with 95% confidence intervals from all sites combined (n = 480). Horizontal line is the 95% line.

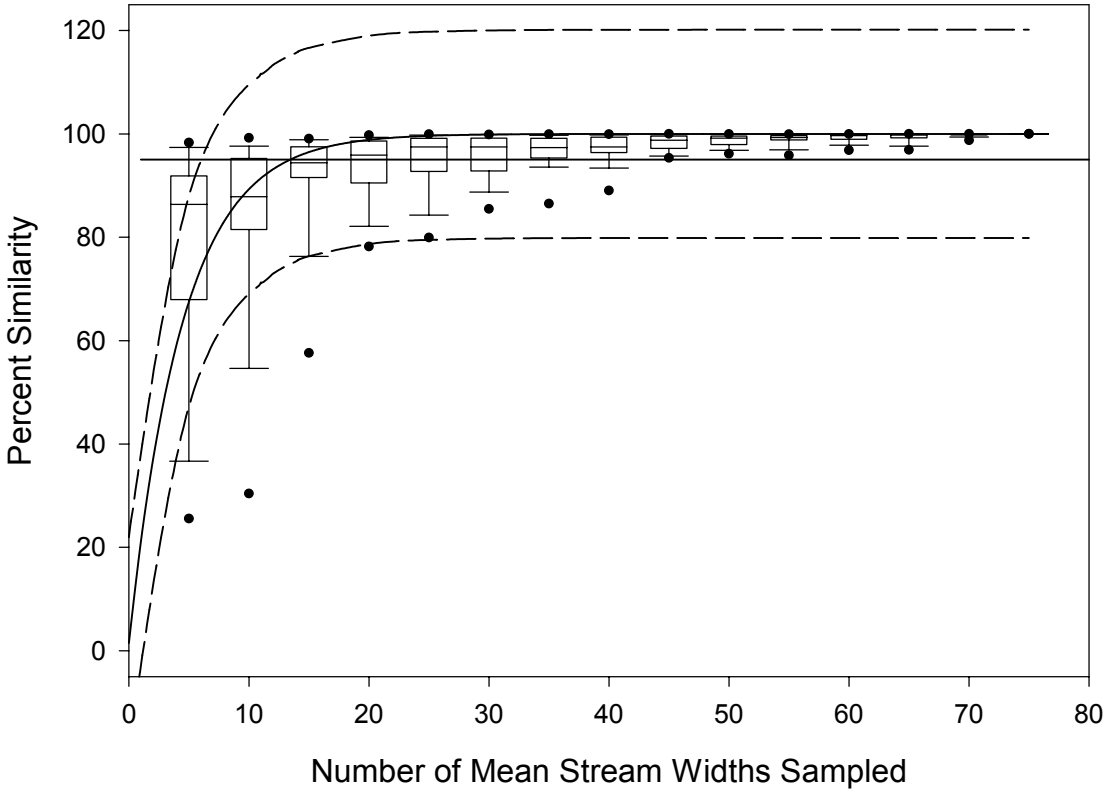


Figure 1.3. Box plots and best-fit line representing the Simplified Morisita's Index of Overlap multiplied by 100 expressing the percent similarity between the number of mean stream widths (MSW) sampled and all segments combined (75 MSW). Middle line in box plots represents the median, box edges are the 25th and 75th percentiles, t-bars are the 10th and 90th percentiles, and dots represent the 5th and 95th percentiles. Predicted line $y = 100.0000-98.5001e^{-0.2220x}$ with 95% confidence intervals from all sites combined (n = 480). Horizontal line is the 95% line.

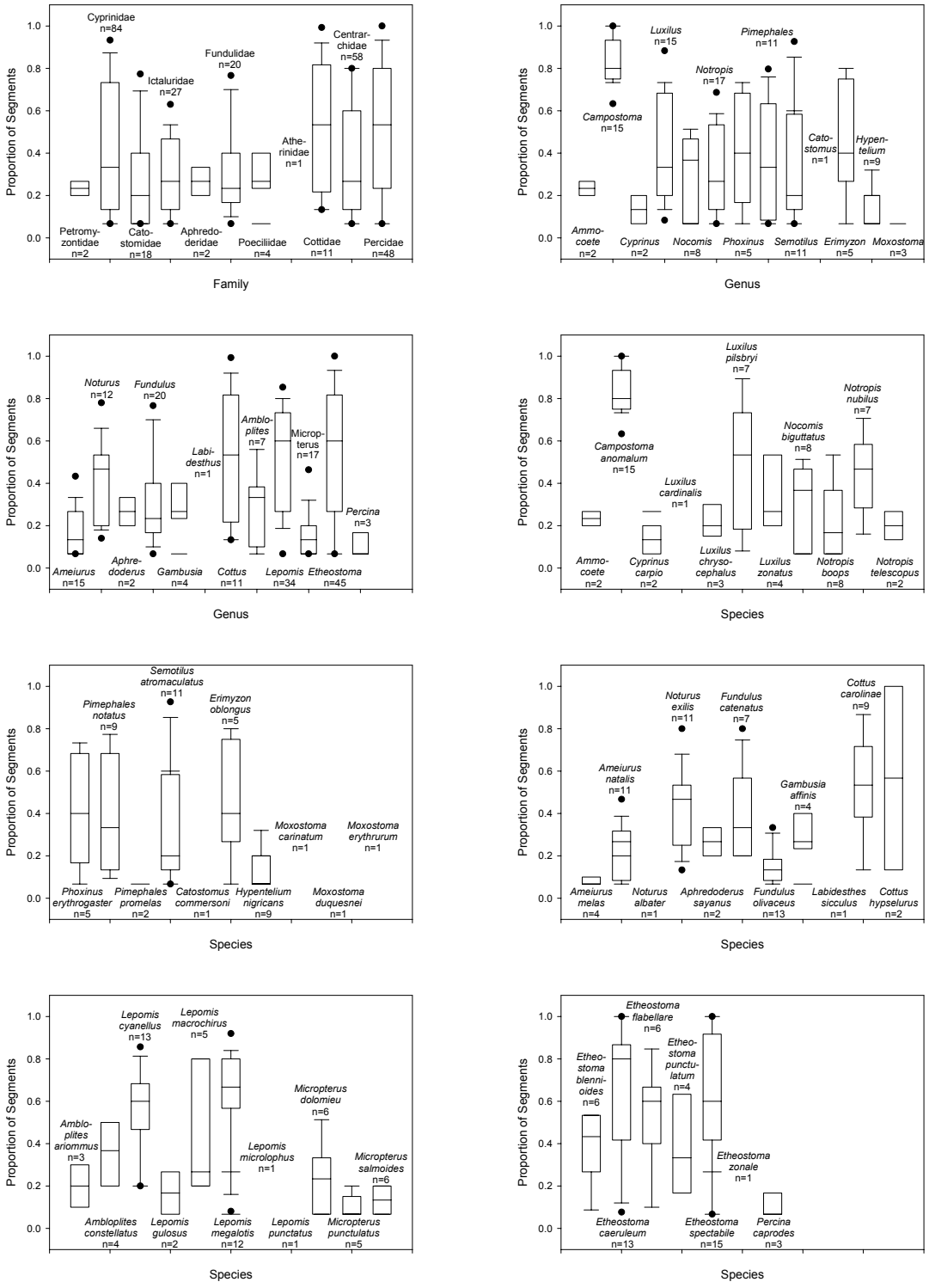


Figure 1.4. Box plots representing the proportion of segments in which each taxa were sampled at each site (see methods for a detailed example). Middle line in box plots represents the median, box edges are the 25th and 75th percentiles, t-bars are the 10th and 90th percentiles, and dots represent the 5th and 95th percentiles.

CHAPTER 2

An Index of Biotic Integrity Developed Using Fish Assemblages for Arkansas' Ozark Highlands Ecoregion

Abstract

The Index of Biotic Integrity (IBI) is a tool used to assess aquatic resource conditions in the U.S. and abroad. We developed an IBI for wadeable streams in the Ozark Highlands ecoregion of Arkansas, USA using an existing fish-collection database and additional data that we collected to augment existing fish collections. We used data from 96 sites to develop and validate the IBI. All 96 fish collections were rarified to simulate fish collections sampled in a stream reach of 51 mean stream widths (MSW) in length, or an equivalent effort. We chose 51 MSW because we previously determined 51 MSW sufficient to estimate fish-assemblage characters pertinent to the IBI in Arkansas' Ozark Highland streams. We classified all fish-collection sites as reference or non-reference based on both subjective and objective information. Nineteen fish collections, 3 reference and 16 non-reference, were removed from IBI development procedures to assess the consistency of reference and non-reference classifications after completing the IBI. We used uni- and bivariate statistics to examine 39 potential IBI metrics. Potential metrics that showed significant differences between reference and non-reference sites, and were not redundant, were included in the IBI. Ten of 39 potential metrics were chosen for the IBI. All IBI metrics were scored from 0 to 10; 0 indicated a strong deviation from reference condition and 10 represented reference condition. IBI scores were calculated to range from 0 to 100; 0 indicated no fish were collected and 100 indicated all metrics reflected reference condition. Trophic metrics contributed most to IBI scores. Sampling effort affected IBI scores, mainly through taxonomic richness metrics. The IBI metrics were most often correlated with nutrients, land use, road density, and sedimentation levels. Index of Biotic Integrity scores calculated for all fish

collections indicate that our fish-collection data is skewed towards good quality sites, which likely represents the current distribution of stream conditions in the Ozark Highlands ecoregion. This skewed distribution may help explain why many non-reference sites were classified as reference sites by the IBI. Based on how the IBI scored sites across the range of different levels of degradation, we feel that our metric selection process and scoring criteria yielded an IBI that can successfully determine stream conditions in Arkansas' wadeable, Ozark Highland streams.

Introduction

In 1972 the U.S. Water Pollution Control Act was implemented to protect U.S. waters. Now known as the Clean Water Act (CWA), this statute mandates state and federal agencies to assess and monitor U.S. surface water trends. Water-quality monitoring programs implemented to meet CWA requirements have been conducted using a variety of chemical and biological methods. Biological endpoints can be advantageous, especially regarding the public's understanding of water-quality goals (Barbour et al. 1999).

Karr (1981) developed the Index of Biotic Integrity (IBI), a multimetric index using fish-assemblage attributes, to assess stream-site quality. The IBI was developed to provide a quick, reliable, and easily understood method to assess stream-site quality (Karr et al. 1986), descriptors not prevalent in previous water-quality monitoring programs (Ward et al. 1986). Since its development, IBI metrics have been modified (see Simon and Lyons (1995) for a summary) for regional (Fausch et al. 1984; Miller et al. 1988), Mexican and European (Lyons et al. 1995; Didier and Kestemont 1996), and specific (Minns et al. 1994; Lyons et al. 1996) applications. The IBI has also been adopted by many states (e.g., Ohio) to assess the status of their waterbodies in compliance with section 305(b) of the CWA.

Successful IBIs require some regional framework for their development, and Omernik's (1987) ecoregions are widely used to define such regions in the conterminous U.S. (Hughes et al. 1987; Angermeier et al. 2000). Distinct fish assemblages occur in the state of Arkansas (Matthews and Robison 1988) and they have been shown to coincide with ecoregions in Arkansas (Keith 1987; Rohm et al. 1987). In fact, an IBI has been

developed using fish assemblages in Arkansas' Ouachita Mountains ecoregion (Hlass et al. 1998). The IBI reflected stream conditions at sites impacted by different timber harvest techniques. Although a multimetric index using macroinvertebrates has been developed for Missouri Ozark streams (Jones et al. 1981), no such index exists to monitor and assess Ozark streams in Arkansas. We chose Arkansas' Ozark Highlands ecoregion as a regional framework to continue IBI development in the state. Because fish have several advantages as biological indicators, and fish-collection data are available through the Arkansas Department of Environmental Quality's (ADEQ) fish-collection database, we used fish assemblages as our biological measure to develop an IBI for the Ozark Highlands ecoregion. Our goal was to use fish assemblages to develop an IBI for Arkansas' wadeable, Ozark Highland streams that will successfully differentiate stream conditions. This goal was achieved through several objectives: 1) identify non-redundant metrics that can differentiate between reference and non-reference sites; 2) determine the percentage agreement between original reference-site classifications; 3) determine relative metric contributions to IBI scores; 4) determine the effects of sampling effort on IBI scores; and 5) determine relations between IBI metrics and physiochemical and land-use variables.

Study Area

The Ozark Highlands ecoregion is located in north-central and north-western Arkansas (Figure 2.1). Mountainous terrain, steep gradients, and fractured limestone geology characterize (Robison and Buchanan 1988) and form the fast-flowing, spring-fed streams of the region. Land use is a major cause of regional water-quality problems.

Many of these problems result from high animal production wastes (e.g., poultry) that have a potential to contaminate regional surface and ground waters because of local geology (ADPCE 1996). The ichthyofauna in this ecoregion is highly diverse, consisting of at least 89 species with cyprinids, percids, and centrarchids contributing most to Ozark streams' relative abundance (Giese et al. 1987).

Methods

Data Collection

We used two sources of fish-assemblage data to develop the Ozark Highlands IBI. One source was the ADEQ's fish-collection database. This database contained 80 Ozark Highland fish collections made across the ecoregion (Figure 2.1) from 1963 to 1999. All but four collections were made after 1982. Most fish collections were made during the summer, but all collections were made when streams represented base flow conditions (W. Keith, ADEQ, pers. comm.). ADEQ personnel sampled varying stream distances among individual streams, using a one-pass electrofishing sample, until they thought that no new fish species were being collected and no new types of habitat were being sampled. This method most likely sampled all or nearly all species present. The second data source was 16 fish collections we made (hereafter referred to as UAPB sites) between June 6 and August 1, 2000, when streams were at or near base flow (Figure 2.1). Thus, the combined data set comprised 96 fish collections. We made our collections to ensure that a wide range of stream-disturbance levels were represented for IBI development and to determine an adequate, one-pass backpack electrofishing sample effort for characterizing fish assemblages (e.g., species richness and assemblage

structure) in wadeable, Ozark Highland streams (Chapter 1). By working with ADEQ on stream sampling, we estimated ADEQ's sampling distance to be approximately 75 MSW when sampling with a backpack electrofisher. In an attempt to replicate the ADEQ's fish sampling protocol we sampled a distance of 75 mean stream widths (MSW) at each site using a backpack electrofisher. Using 15 of our 16 fish collections we determined that 51 MSWs need sampling to characterize both species richness and assemblage structure of fish assemblages in wadeable Ozark Highland streams.

Some ADEQ fish samples were collected using a barge electrofisher. Barge electrofishing was used in larger streams where it was considered more effective than backpack electrofishing. Streams were sampled extensively by barge electrofishing until ADEQ personnel thought no new species were being collected and no new habitats were being sampled. Relations between sampling effort and fish-assemblage characters were not determined using this sampling gear. Other researchers (e.g., Paller (1995) and Patton et al. (2000)) found a lower number of MSWs need sampling in larger streams, while Lyons (1992) suggested there "could be" a maximum sampling distance in Wisconsin streams. Although we found that MSW was a good scaling factor for the streams we sampled, we sampled only smaller streams using a backpack electrofisher. A potentially lower number of MSWs, but an equivalent relative effort, may need sampling in larger streams that require sampling by barge electrofishing.

Because all ADEQ and UAPB fish samples were collected from sites 75 MSW (or approximately that distance) in length using backpack electrofishing, or an equivalent effort in larger streams requiring barge electrofishing, all fish collections were rarified by randomly removing 32% (i.e., (1 - (51 MSW / 75 MSW)) x 100) of the fish from each

collection. This rarification created fish-collection data from each site that represented samples collected in a stream reach of 51 MSWs in smaller streams, and an equivalent effort in larger streams.

Site Classifications

We classified all 96 fish-collection sites as reference or non-reference sites. Classifications were based on best professional judgment and land-use patterns. The ADEQ biologists involved in sampling fishes throughout Arkansas classified the ADEQ fish-collection sites as reference or non-reference sites on two different occasions. Classifications were based on best professional judgement. Sites classified as reference sites on both occasions continued to be considered for reference-site classification. To classify the UAPB sites, each author independently classified all 16 sites as reference or non-reference sites. Sites classified as reference sites by both authors continued to be considered for reference site classification. For IBI development, we classified sites as reference sites when they were classified as reference sites by ADEQ personnel or us, and had at least 75% forested watersheds. Therefore, we had dual criteria for reference-site classifications (hereafter referred to as "dual criteria"). The exceptions to the duel criteria rule were fish collections from 1963. ADEQ personnel classified these sites as reference sites, but appropriate temporal land-use data were not available to determine land use percentages. We assumed these sites reflected reference condition, and classified them as reference sites. Each site's contributing watershed was delineated using ArcView® software, and we determined watershed land use using the Arkansas Gap Analysis (AR-GAP) land-cover data layer (Smith et al. 1998; Weih 2001). Landsat

Thematic Mapper (TM) data, mostly from 1992, was used as the base data layer to develop AR-GAP land-cover classifications. The final AR-GAP data layer includes 36 land-cover classes with a 100 ha resolution. We grouped the AR-GAP land-cover classifications into agriculture, forested, and urban land-use classes.

We omitted approximately 20% of the total sites from use in metric selection and scoring procedures to later assess the consistency of IBI site classifications (i.e., reference vs. non-reference). To ensure that the removed data set included a range of stream sizes, we stratified all sites into watershed size-groups of 0 – 100 km^2, >100 – 300 km^2, and >300 km^2. We then randomly removed 20% of the sites from each watershed size-group in the data set. One reference site was removed from each size-group. Nineteen sites were removed from the complete data set.

Candidate IBI Metrics

We selected candidate metrics for the IBI from a variety of sources (e.g., Simon and Lyons (1995)). We also considered some novel metrics that seemed potentially useful for the Ozark Highlands ecoregion. Only metrics that had non-zero values for most sites were considered as candidate IBI metrics. As recommended by Karr et al. (1986), juvenile fish were excluded from all fish collections. An exception was lamprey ammocoetes (*Ichthyomyzon* and *Lampetra* juveniles). Juvenile lampreys are easily identified as such and they are potentially important in determining a stream site's condition because they indicate a migratory adult's presence in a stream. We classified all fish species into taxonomic, functional, reproductive, and trophic categories. Classifications were assigned by using Ohio EPA (1987), Robison and Buchanan (1988),

Etnier and Starnes (1993), Jenkins and Burkhead (1993), Smogor (1996), Pflieger (1997), Smogor and Angermeier (1999), the ADEQ's fish classifications, and professional judgement. Individual metrics were developed in the categories of taxonomic richness and composition, trophic composition, reproductive richness and composition, and fish abundance and condition. Taxonomic richness and composition metrics were developed with their respective taxa. We used Smogor and Angermeier's (1999) trophic classification procedure where food types were grouped into detritus (DET), algae-vascular plants (AH), invertebrates (INV), and fish (P). All fish species were determined to eat one or any combination of these food groups and were classified as: DET/AH = 1, DET/AH/INV = 2, DET/AH/INV/P = 3, AH/INV = 4, AH/INV/P = 5, INV = 6, INV/P = 7, and P = 8. We classified fish reproductive behaviors similar to Smogor and Angermeier (1999). These classifications were based on whether or not they require mineral substrates exclusively, prepare the substrate prior to spawning, and provide parental care. We considered the following combinations of these reproductive strategies for potential metrics: simple (no site preparation or parental care), lithophilic (mineral substrate) spawners = 1, mineral, site-prep spawners = 2, mineral, parental-care spawners = 3, mineral, site-prep, parental-care spawners = 4, miscellaneous, site-prep spawners = 5, miscellaneous, parental-care spawners = 6, miscellaneous, site-prep, parental-care spawners = 7, and simple, miscellaneous spawners = 8. Nest associates were considered site-prep spawners because they required a modified substrate to spawn. Non-taxonomic fish classifications are listed in Table 2.1. The candidate metrics we tested and their associated acronyms (listed in Table 2.2) are listed and described below:

Species Richness and Composition

Total Species (TSPECI): Indicates the total number of species in a sample site. Higher numbers of species are expected at reference sites.

Total Darter Species (TDARTE): Indicates the total number of darter species in a sample site. This metric includes both *Etheostoma* and *Percina* genera from the family Percidae. Higher numbers of darter species are expected at reference sites.

Total Darter and Sculpin Species (TDARSC): Indicates the total number of darter (*Etheostoma* and *Percina spp.*) and sculpin (*Cottus spp*.) species in a sample site. Higher numbers of darter and sculpin species are expected at reference sites.

Total Darter, Sculpin, and Madtom Species (TDASCM): Indicates the total number of darter (*Ethoestoma* and *Percina spp.*), sculpin (*Cottus spp.*), and madtom (*Noturus spp.*) species in a sample site. Higher numbers of darter, sculpin, and madtom species are expected at reference sites.

Total Sunfish Species (TSUNFI): Indicates the total number of sunfish species (*Ambloplites* and *Lepomis spp.*) in a sample site. Higher numbers of sunfish species are expected at reference sites.

Total Sucker Species (TSUCKE): Indicates the total number of sucker species (Family: Catostomidae) in a sample site. Higher numbers of sucker species are expected at reference sites.

Total Intolerant Species (TINTOL): Indicates the total number of intolerant species in a sample site. These classifications are from the ADEQ fish classifications and are based on a survey of various fisheries experts in Arkansas concerning the

sensitivity of Arkansas' fish species. Higher numbers of intolerant species are expected at reference sites.

Total Benthic Species (TBENTH): Indicates the total number of benthic species in a sample site. Benthic species primarily reside on the stream bottom. Higher numbers of benthic species are expected at reference sites.

Total Cyprinid Species (TCYPRI): Indicates the total number of species from the family Cyprinidae in a sample site. Higher numbers of Cyprinid species are expected at reference sites.

Total Shiner Species (TSHINE): Indicates the total number of shiner species in a sample site. This metric includes *Cyprinella*, *Clinostomus*, *Lythrurus*, *Luxilus*, and *Notropis spp*., but excludes the golden shiner (*Notemigonus crysoleucas*). Higher numbers of shiner species are expected at reference sites.

Percentage of Individuals as Green Sunfish (PGRESF): Indicates the percentage of individuals in the fish assemblage that is green sunfish (*Lepomis cyanellus*). Higher percentages of green sunfish are expected to occur at non-reference sites.

Percent of Individuals as Green Sunfish and Yellow Bullhead (PGSFYB): Indicates the percentage of individuals in the fish assemblage that is green sunfish (*Lepomis cyanellus*) and yellow bullhead (*Ameiurus natalis*). Higher percentages of green sunfish and yellow bullheads are expected to occur at non-reference sites.

Percent of Individuals as Green Sunfish, Bluegill, Yellow Bullhead, and Channel Catfish (PGBYCC): Indicates the percentage of individuals in the fish assemblage that is green sunfish (*Lepomis cyanellus*), bluegill (*Lepomis macrochirus*), yellow bullhead (*Ameiurus natalis*), and channel catfish (*Ictalurus punctatus*). Higher

percentages of green sunfish, bluegill, yellow bullhead, and channel catfish are expected to occur at non-reference sites.

Percentage of Individuals as Stonerollers (PSTONE): Indicates the percentage of individuals in the fish assemblage that is stonerollers (*Campostoma spp.*). Higher percentages of stonerollers are expected to occur at non-reference sites.

Percentage of Individuals as Intolerant Species (PINTOL): Indicates the percentage of individuals in the fish assemblage that is intolerant species. These classifications are from the ADEQ fish classifications and are based on a survey of various fisheries experts in Arkansas concerning the sensitivity of Arkansas' fish species. Higher percentages of intolerant species are expected to occur at reference sites.

Percent of Individuals as Bethics (PBENTH): Indicates the percentage of individuals in the fish assemblage that are benthic species. Benthic species primarily reside on the stream bottom. Higher percentages of benthic species are expected to occur at reference sites.

Percent of Individuals as Darters and Sculpins (PDARSC): Indicates the percentage of individuals in the fish assemblage that is darter (*Etheostoma and Percina spp.*) and sculpin (*Cottus spp.*) species. Higher percentages of darter and sculpin species are expected to occur at reference sites.

Percent of Individuals as Darters, Sculpins, and Madtoms (PDASCM): Indicates the percentage of individuals in the fish assemblage that are darter (*Etheostoma and Percina spp.*), sculpin (*Cottus spp.*), or madtom (*Noturus spp.*) species. Higher percentages of darter, sculpin, and madtom species are expected to occur at reference sites.

Percent of Individuals as Cyprinids (PCYPRI): Indicates the percentage of individuals in the fish assemblage that is in the family Cyprinidae. Higher percentages of Cyprinid species are expected to occur at reference sites.

Percent of Individuals as Shiners (PSHINE): Indicates the percentage of individuals in the fish assemblage that is shiners. This metric includes *Cyprinella*, *Clinostomus*, *Lythrurus*, *Luxilus*, and *Notropis spp*., but excludes the golden shiner (*Notemigonus crysoleucas*). Higher percentages of shiner species are expected to occur at reference sites.

Reproductive Composition

Total Mineral Spawning Species (TMINSP): Indicates the total number of mineral (i.e. lithophilic) spawning species in a sample site. These species have a reproductive classification of 1, 2, 3, or 4. Higher numbers of mineral spawning species are expected at reference sites.

Total Simple, Lithophilic Spawning Species (TSIMLI): Indicates the total number of simple, lithophilic spawning species in a sample site. These species have a reproductive classification of 1. Higher numbers of simple, lithophilic spawning species are expected at reference sites.

Percent of Individuals as Simple, Lithophilic Spawners (PSIMLI): Indicates the percentage of individuals in the fish assemblage that is simple, lithophilic spawners. These species have a reproductive classification of 1. Higher percentages of simple, lithophilic spawning species are expected to occur at reference sites.

Percent of Individuals as Mineral, Site-Prep Spawners (PMINSP): Indicates the percentage of individuals in the fish assemblage that prepare spawning sites on mineral substrates. These species have a reproductive classification of 2. Higher percentages of mineral, site-prep spawners are expected to occur at reference sites.

Percent of Individuals as Simple, Miscellaneous Spawners (PSIMPX): Indicates the percentage of individuals in the fish assemblage that is simple spawners that use miscellaneous substrates. These species have a reproductive classification of 8. Higher percentages of simple, miscellaneous spawning species are expected to occur at non-reference sites.

Percent of Individuals as Miscellaneous, Site-Prep, Parental-Care Spawners (PXSPPC): Indicates the percentage of individuals in the fish assemblage that prepare miscellaneous substrates for spawning and provide parental care to their offspring. Species included in this metric have reproductive classification of 7. Higher percentages of miscellaneous, site prep, parental care spawners are expected to occur at non-reference sites.

Trophic Composition

Total Generalist Species (TGENER): Indicates the total number of generalist species in a sample site. Generalists feed on more than two food types and have trophic classifications of 2, 3, or 5. Higher numbers of generalist species are expected at non-reference sites.

Percent of Individuals as Generalists (PGENER): Indicates the percentage of individuals in the fish assemblage that is generalist feeders. Generalists feed on more than two food types and have trophic classifications of 2, 3, or 5. Higher percentages of generalists are expected to occur at non-reference sites.

Percentage of Individuals as Top Carnivores (PTOPCA): Indicates the percentage of individuals in the fish assemblage that is top-level carnivores. Top-level carnivores are species that are capable of having primarily piscivorous diets as adults. Higher percentages of top-level carnivores are expected to occur at reference sites.

Percent of Individuals as Algivorous/Herbivorous, Invertivorous, and Piscivorous (PAHINP): Indicates the percentage of individuals in the fish assemblage that consume algae, vascular plants, invertebrates, and fish. Species included in this metric have a trophic classification of 5. Higher percentages of AH/INV/P species are expected to occur at non-reference sites.

Percent of Individuals as Invertivores and Piscivores (PINVPI): Indicates the percentage of individuals in the fish assemblage that consume invertebrates and fish. Species included in this metric have a trophic classification of 7. Higher percentages of INV/P species are expected to occur at reference sites.

Percent of Individuals as Invertivores (PINVER): Indicates the percentage of individuals in the fish assemblage that consume invertebrates. Species included in this metric have a trophic classification of 6. Higher percentages of invertivorous species are expected to occur at reference sites.

Percent of Individuals as Detritivores and Algivore/Herbivores (PDETAH): Indicates the percentage of individuals in the fish assemblage that consume detritus, algae, and vascular plants. Species included in this metric have a trophic classification of 1. Higher percentages of D/AH species are expected to occur at non-reference sites.

Percent of Individuals as Algivore/Herbivores and Invertivores (PAHINV): Indicates the percentage of individuals in the fish assemblage that consume algae, vascular plants, and invertebrates. Species included in this metric have a trophic classification of 4. Higher percentages of AH/INV species are expected to occur at non-reference sites.

Fish Abundance and Condition

Individuals per Second (INDSEC): Indicates the number of individuals collected per second of electrofishing. Higher numbers of individuals per second are expected to occur at reference sites.

Percent of Individuals with Anomalies (PANOMA): Indicates the percentage of individuals in the fish assemblage that has external anomalies. Anomalies considered are: deformities, lesions, tumors, fish erosion, all grubs (excluding black-spot disease (*Neascus spp.*)), and leeches. Higher percentages of fish with anomalies are expected at non-reference sites.

Percent of Individuals with Black Spot (PBLKSP): Indicates the percentage of individuals in the fish assemblage that has at least one visible black spot cyst (*Neascus spp.*). Higher percentages of fish with at least one visible black spot cyst are expected to occur at non-reference sites.

Percent of Individuals with Black Spot 3+ (PBLK3+): Indicates the percentage of individuals in the fish assemblage that has at least three visible black spot cysts (*Neascus spp.*). Higher percentages of fish with at least three visible black spot cysts are expected to occur at non-reference sites.

Percent of Individuals with Black Spot or Anomaly (PBLKAN): Indicates the percentage of individuals in the fish assemblage that has any anomaly, including visible black spot cysts (*Neascus spp.*). Anomalies considered are: deformities, lesions, tumors, fish erosion, all grubs (including black spot cysts), and leeches. Higher percentages of fish with visible black spot cysts or an anomaly are expected at non-reference sites.

Metric Selection

We used two criteria to choose the final IBI metrics from the candidate metrics. The first criterion was a metric's ability to differentiate between reference and non-reference sites. For the metrics that met the first criterion we tested for metric redundancy, our second criterion.

We tested each candidate metric to determine if a difference existed between dual-criteria classified reference and non-reference sites. We performed two-tailed t-tests assuming equal or unequal variances, depending on F-test results, to determine if differences existed between metric values of reference and non-reference sites ($\alpha = 0.10$). We chose an alpha level of 0.10 because of the natural variability Ozark Highland streams and also because stream size may cause overlap between reference and non-reference values for certain metrics. All metrics consisting of proportional data were

arcsin transformed to meet normality assumptions (Zar 1999). Metrics showing significant differences between reference and non-reference data continued to be considered as candidate metrics.

We tested for metric redundancy by conducting simple linear correlations between all combinations of metrics. For each metric we used reference and non-reference data from all sites (Barbour et al. 1992). High and low correlation coefficients ($r > 0.90$ or $r < -0.90$) indicated redundant metrics (Angermeier et al. 2000). Some metrics were expected to be highly correlated because they contain similar taxa (e.g., TDARSC and TDASCM), or they may have similar functional characteristics (e.g., TDASCM and TBENTH). The metric from each redundant metric group that showed the lowest probability of a type I error from the previous t-test was included in the IBI.

We determined which IBI metrics were significantly ($\alpha = 0.05$) affected by stream size by conducting linear regressions between watershed size (independent variable) and the selected IBI metrics (dependent variable). Each analysis included only reference-site data (i.e., $n = 10$) in order to eliminate any confounding effects resulting from stream condition. This analysis allowed us to determine which metrics require scoring criteria that adjusts for stream size. We assumed a linear relationship between watershed size and those metrics affected by watershed size (Figure 2.2).

Metric Scoring

We scored metrics similar to Minns et al. (1994) by developing threshold values used to standardize metric scoring criteria from 0 – 10. We slightly deviated from the

procedures of Minns et al. (1994) when scoring metrics that were affected by watershed size.

We scored all metrics using upper and lower threshold limits from each metric's data range. Determining threshold limits depended on a metric's relation with stream-site quality. When raw metric values for reference sites were higher than non-reference sites (positively related) the 50th percentiles of the reference-site data for those metrics were used as the upper threshold. Zero was used as the lower threshold. Metrics with lower values for reference sites than non-reference sites (negatively related) had their lower thresholds set at the 50th percentiles of that metric's reference-site data. The upper threshold was set at the 95th percentile of the non-reference site data. We used the 95th percentile as the upper threshold for negatively related metrics because our data set had a limited number of poor-quality sites and a high proportion of high-quality sites. All metrics affected by stream size were positively related with stream quality and were scored using two different methods. One method was used for sites with watersheds between 0 and 800 km^2. This method used upper threshold limits determined from best-fit linear-regression lines of watershed size versus metric values. Only reference-site data were used. Thus, watershed size was the independent variable used to determine each site's upper threshold limit. The second method was used on sites with watersheds greater than 800 km^2. These sites all had the same upper threshold as a site with a watershed size of 800 km^2 (Figure 2.3). Lower thresholds for affected metrics were set at zero for all watershed sizes.

After determining threshold limits, we adjusted each metric to score from 0 (very poor condition) to 10 (reference condition). All metric scores were derived using the equation

$$MS = A + B*(MR)$$

where *MS* = Metric Score, *MR* = Raw Metric Value, *A* = the y-intercept in the regression of *MS* versus *MR*, *B* = the slope in the regression of *MS* versus *MR*. All *MS* (dependent) versus *MR* (independent) regressions included predetermined points; raw metric thresholds received predetermined metric scores (0 or 10) depending on each metric's relation with stream condition. For example, the metric PINTOL has an upper threshold of 57.10 and a lower threshold of 0.00. Thus, the points used in the regression to determine the regression coefficients were (57.10, 10.00) and (0.00, 0.00). In the above metric scoring equation the following conditions must be in place

$$\text{If } MR < LT, \text{ then } MR = LT$$

$$\text{If } MR > UT, \text{ then } MR = UT$$

where *LT* equals the lower threshold limit and *UT* equals the upper threshold limit. Thus, the equation calculates a metric score from the raw metric value and upper and lower threshold limits. Threshold limits define the maximum and minimum values a raw metric value may have when included in the equation. Raw metric values above the upper threshold limit and below the lower threshold limit take the value of each respective threshold. All taxonomic-richness metrics affected by watershed size had their threshold limit values rounded to the nearest integer. Therefore, all metrics were scored from 0 to 10.

IBI Scoring and Qualitative Classifications

Individual metric scores were used to calculate IBI scores ranging from 0 to 100. IBI scores are calculated as follows

$$IBI = \frac{\left(\sum_{i=1}^{n} MS_i\right) \cdot 10}{n}$$

where *IBI* = IBI score, *MS* = metric score of the *i*th metric, and *n* = the number of metrics. The scoring ranges for the qualitative integrity classifications are 1 - <20 (very poor), 20 - <40 (poor), 40 - <60 (fair), 60 - <80 (good), and 80-100 (reference). An IBI score of zero should be assigned if no fish were collected. Qualitative classifications can be used to interpret IBI scores for discerning stream-site quality.

Upon completion of the IBI, distributions of original reference-site IBI scores were used to set a threshold for classifying stream sites as impaired. The 25th percentile of reference site scores was used in an example by Barbour et al. (1999). We used the 25th percentile as a starting point for determining a threshold IBI score, but we also used a frequency histogram of reference site IBI scores to detect gaps in reference-site IBI scores. These two methods yielded our IBI score discriminating impaired versus unimpaired stream sites.

IBI Precision

Karr (1981) originally recommended IBI scoring ranges between the qualitative site-classification scoring ranges to be used for professional judgement of stream-site quality. These ranges are recommended because inherent variability in IBI scores, due to the potential dynamic nature of stream-fish assemblage structure and/or sampling error,

may cause a stream site that scores close to a qualitative classification threshold score to be mis-classified. For example, a stream site with an IBI score of 59 may be classified as being in "fair" condition when it is actually in "good" condition.

To recommend scoring ranges between our recommended qualitative scoring ranges we estimated the IBI's within-site precision using a bootstrap technique similar to Fore et al. (1994). The fifteen segmented stream sites (UAPB sites; Chapter 1) were used for the analysis. Although fifteen segments were sampled in the field, only 10 (equal to approximately 51 MSWs) were recommended for IBI calculations (Chapter 1). Therefore, for each site we randomly selected, without replacement, 10 stream segments 25 times per stream site. Hereafter, these samples are referred to as bootstrap samples. We then calculated IBI scores using fishes from each bootstrap sample. The PBLKAN metric scores were not included in IBI scores because segment data for that metric were not available. Ranges and 95% confidence intervals were calculated for IBI scores computed for each stream site. Ranges and confidence-interval lengths were used to determine an IBI scoring range that can be used for professional judgement of stream-site quality when a stream site receives an IBI score near or at a classification threshold.

Relations between confidence-interval lengths and mean IBI scores of bootstrap samples, total number of individuals, and species richness of all 15 sites were assessed. Relations were assessed using simple linear correlations at $\alpha = 0.05$. Total number of individuals and species richness of all 15 sample segments of each sample site were used for correlation analyses.

Classification Consistency

We used the 19 randomly removed sites to assess the consistency of reference/non-reference classifications given by the IBI, ADEQ personnel, and dual-criteria site classifications. To do this we determined the percent agreement between site classifications based on the IBI scores and both those based on the professional judgement of the ADEQ personnel and the dual criteria. We also performed the same analysis using all fish-collection sites (i.e., including the 19 sites that were removed for IBI development). For this analysis we used an IBI score of 80 or higher to represent reference condition.

Metric Contributions

We used two methods to evaluate each metric's relative contribution to IBI scores in all qualitative classifications combined and for IBI scores in each exclusive qualitative classification. Metric contributions to IBI scores for all qualitative classifications were measured using data from all sites. Metric contributions to IBI scores in each qualitative classification were measured using data from sites with IBI scores in each respective classification. The first method we used to assess metric contributions was based on a simple linear correlation coefficient. This technique correlates a metric score with an IBI score calculated without the specific metric under investigation (Hughes et al. 1998). Because we previously screened the IBI metrics for signals of reference condition, we interpreted a lower correlation coefficient as reflecting a higher metric contribution to the IBI. The second method measured the variance of the differences between IBI scores computed with and without an individual metric (Minns et al. 1994). A higher variance

reflects a higher metric contribution to the IBI. In order to evaluate the relative contribution of each IBI metric we ranked each metric according to its contribution measure for each method (i.e., correlation coefficient and variance). We then ranked the sums of both ranks for each metric to determine its relative influence on IBI scores for all qualitative classifications combined and individual qualitative classifications. No sites contained IBI scores indicating very poor stream condition (i.e., a score ranging from >0 - <20). Metric contribution evaluations were not completed for the PBLKAN metric in the 20 - <40 (i.e., poor) scoring range because only one site in that scoring range contained data for that metric.

Sampling Effort and IBI Relations

We used 15 of our sample sites to assess the effects of sampling effort on IBI scores. These sites were segmented into 15 contiguous segments; each segment was five mean stream widths (MSW) in length (see Chapter 1 for details). Fishes were sampled in each segment by backpack electrofishing. Index of Biotic Integrity scores were calculated using fishes from accumulating segments in both upstream and downstream directions at each site (n = 30). Therefore, an IBI score was calculated every five MSWs in each direction. Significant effects of sampling effort on IBI scores and individual metric scores were detected using repeated-measures ANOVA and Tukey test ($\alpha = 0.05$) on IBI scores and metric scores calculated every five MSWs. Mean differences in IBI scores, and also metric scores, between adjacent segments (i.e., ((IBI score for n MSWs + 5) – IBI score for n MSWs)) were plotted to assess scoring and variance trends between sampling distances. For example, the IBI difference value at 35 MSWs is the mean

difference (± SE) of IBI scores at 40 MSWs minus the IBI scores at 35 MSWs, for all sites. We also determined the percentage of sites with different qualitative classifications from their classifications at 50 MSWs, when they were sampled at distances other than 50 MSWs (also in upstream and downstream directions (n = 30)). Mean differences (± SD) between IBI scores calculated every 5 MSWs and IBI scores at 50 MSWs were plotted to determine differences in IBI scores at different sampling distances. We compared site classifications and IBI scores to those at 50 MSWs because the IBI was calibrated for a sampling effort of 51 MSWs. These analyses convey possible effects of sampling a distance other than the sampling effort for which the IBI was calibrated, and therefore, should offer insight into the robustness of the IBI to under- and over-sampling bias.

Environment and IBI Relations

We also determined relations between raw values of our IBI metrics and selected physiochemical and land-use variables at each site. A Spearman's rank correlation was used to relate IBI metrics to the selected variables. Selected variables included: percent watershed as forested land use, percent forested land use in a 100 m stream buffer, percent watershed as agricultural land use, percent agricultural land use in a 100 m stream buffer, percent watershed as urban land use, percent urban land use in a 100 m stream buffer, watershed road density (km/km^2), road density in a 100 m stream buffer (km/km^2), visually assessed sedimentation levels (ranging from 0 (low) to 15 (high)), riparian zone width, and various water-quality variables.

Land-use variables, as determined using the AR-GAP data layer used in selecting reference sites, were correlated with IBI metrics. Riparian zone width was measured from digital orthographic quadrangles (DOQs; 1-meter resolution) and GIS software. Sample sites were identified and riparian zone width on the left and right banks were measured perpendicular to the stream channel at 10 equidistant transects covering the sample site. Therefore, 20 riparian-zone-width measurements were taken. The maximum riparian-zone-width measurement was 30 m; the median riparian-zone width was used in correlation analyses.

Water-quality variables used in the correlations analyses included: aluminum (µg/l), arsenic (µg/l), barium (µg/l), boron (µg/l), cadmium (µg/l), calcium (mg/l), chromium (µg/l), cobalt (µg/l), copper (µg/l), iron (µg/l), magnesium (mg/l), manganese (µg/l), nickel (µg/l), potassium (mg/l), sodium (µg/l), vanadium (µg/l), zinc (µg/l), hardness (mg/l), silicate (mg/l), dissolved oxygen (mg/l), pH, water temperature (C), bromide (mg/l), fluoride (mg/l), sulfates (mg/l), ammonia (mg/l), chlorides (mg/l), nitrates (mg/l), ortho-phosphates (mg/l), total phosphorous (mg/l), total kjeldahl nitrogen (mg/l), total organic carbon (mg/l), biological oxygen demand (mg/l), turbidity (mg/l), total suspended solids (mg/l), and total dissolved solids (mg/l). Water-quality data for fish-collection sites were taken from various Arkansas Department of Environmental Quality (formerly the Arkansas Department of Pollution Control & Ecology (ADPC&E)) reports and water samples from our sample sites, which were processed by the ADEQ's water-quality laboratory. Various editions of Standard Methods for Examination of Water and Wastewater and the ADEQ's Quality Assurance Project Plan for Ambient

Water Quality and Compliance Monitoring, 1995, methods were used for sample collection and processing (ADPCE 1993; ADPCE 1995).

Relations among physiochemical variables were assessed using principal components analysis (PCA) with a varimax rotation. Ordination analyses are improved by removing outliers (Gauch 1982), therefore, we removed all outliers from each physiochemical variable. Also, we only included variables containing at least 48 observations. Percent agriculture land-use was not included because proportional variables summing to unity can confound analyses. Percent agriculture land-use was most often a direct reflection of percent forested land-use, because urban land-use is minimal in most Ozark Highland watersheds. In an effort to meet the PCA assumptions of normally distributed data (Gauch 1982) and independent means and variances, proportional data were arcsin transformed and all other data were square-root transformed (Sokal and Rohlf 1995). We were not concerned about the PCA assumption of orthogonal variables because the analysis was used to elucidate relations among variables and not for hypothesis testing. For the descriptive purpose of this analysis, the PCA assumptions do not need to be met completely (Gauch 1982). Further, inflated PCA axis scores resulting from highly correlated variables would enhance the detection of those correlated variables and determine which variables account for most of the model's variance.

Results

Fish Collections

Our fish-collection database comprised 96 fish collections. The random removal of 32% of the individuals from each fish collection removed at least one species from most collections. The average number of species removed per collection was 1.31 (SD = 1.22; range = 0 - 5).

Site Classifications

Our reference site criteria yielded 13 reference sites. The randomly removed sites included 11 sites in the 0 – 100 km^2 group, three in the >100 – 300 km^2 group, and five in the >300 km^2 group, for a total of 19 sites.

Metric Selection and Scoring

Metrics showing a significant difference between reference and non-reference sites included: PAHINP, PBLKAN, PBLKSP, PGBYCC, PGRESF, PGSFYB, PINTOL, PINVER, PSTONE, PTOPCA, PXSPPC, TBENTH, TDARSC, TDARTE, TDASCM, TINTOL, TMINSP, and TSPECI (Table 2.3). Correlation coefficients indicated that 17 pairs of metrics were highly correlated, indicating metric redundancy (Table 2.4). After selecting one metric from all redundant groups (except TSPECI and TMINSP were both selected and redundant), metrics selected for the IBI included: PAHINP, PBLKAN, PGBYCC, PINTOL, PINVER, PSTONE, PTOPCA, TDASCM, TMINSP, and TSPECI. We included the redundant metrics TSPECI and TMINSP because TSPECI has been widely used and is believed to be a very useful metric, and TMINSP represented the only

reproductive metric remaining as a candidate metric. Therefore, we thought including both metrics would benefit the IBI. Of the selected metrics TSPECI, TDASCM, and TMINSP were significantly influenced by watershed size as indicated by linear regression analyses (Table 2.5). Table 2.6 lists scoring criteria for each metric.

Qualitative Classifications

Our population of reference-site IBI scores yielded a 25th percentile of 85.3. The frequency histogram of reference-site IBI scores showed a gap between IBI scores of 86 and 90 (Figure 2.4). Therefore, an IBI score of 85 was chosen as the threshold for determining a site's impairment status. An IBI score greater than 85 indicates an unimpaired site, while an IBI score of 85 or lower indicates impairment. Sixty-three total sites were classified as impaired, while 33 were considered unimpaired. Four sites originally classified as reference sites by the dual criteria were considered impaired, while nine original reference sites were considered unimpaired.

IBI Precision

Ninety-five percent confidence-interval lengths of IBI scores for the randomly selected segment samples for each site averaged 1.9 units (SD = 0.62). The minimum confidence interval was 0.7 units, while the maximum was 3.2 units. Index of Biotic Integrity score ranges averaged 9.0 units (SD = 2.89). The minimum range was 4.0 units; the maximum range was 15.9 units. We used the rounded-up value of the maximum confidence-interval length as the range for professional judgement of qualitative

classifications at or near classification thresholds. This range of 4 IBI units should be centered on each respective threshold.

Correlations indicated no significant relations between confidence-interval lengths and mean IBI score ($r = -0.392$; p-value = 0.447), total number of individuals ($r = 0.104$; p-value = 0.712), and species richness ($r = 0.086$; p-value 0.762) at $\alpha = 0.05$.

Classification Consistency

Comparing qualitative classifications of removed sites resulted in a 43.8% agreement between qualitative classifications generated by the IBI and the ADEQ site classifications. For removed sites, the percent agreement between qualitative classifications generated by the IBI and dual criteria was 47.4%. The IBI and ADEQ site classifications for all sites had a percent agreement of 67.5%, while IBI and the dual criteria site classifications for all sites agreed 60.4% of the time (Table 2.7). All but one of the disagreeing classifications resulted from the IBI classifying sites as reference sites when they were originally pre-classified as non-reference sites.

Metric Contributions

Trophic metrics appeared to be the most influential on IBI scores. The relative contribution of each metric to the IBI ranked as follows: PINVER (1.5), PAHINP (1.5), PTOPCA (3), PBLKAN (4.5), PSTONE (4.5), PGBYCC (6), PINTOL (7), TSPECI (8.5), TDASCM (8.5), and TMINSP (10) (Table 2.8). Metric contributions to IBI scores in qualitative scoring ranges are listed in Table 2.8.

Sampling Effort and IBI Relations

Repeated-measures ANOVA results indicated sampling effort affected IBI scores and the PINVER, PTOPCA, TDASCM, TMINSP, and TSPECI metrics ($p < 0.05$). Multiple comparisons revealed that the effect of sampling effort on IBI and metric scores decreases with increased sampling effort (Figure 2.5; Figure 2.6). The percent of sites having different qualitative site classifications, determined every five MSWs, as compared to the qualitative site classifications at 50 MSWs are: 5 MSWs = 66.7%; 10 MSWs = 56.7%; 15 MSWs = 33.3%; 20 MSWs = 16.7%; 25 MSWs = 13.3%; 30 MSWs = 6.7%; 35 MSWs = 10.0%; 40 MSWs = 10.0%; 45 MSWs = 10.0; 55 MSWs = 0.0%; 60 MSWs = 10.0%; 65 MSWs = 13.3%; 70 MSWs = 10.0; 75 MSWs = 13.3% (Figure 2.7).

Environment and IBI Relations

Spearman's rank correlations indicated that 15 physiochemical and land-use variables were significantly correlated with PINVER, 14 variables were significantly correlated with PGBYCC and TSPECI, 13 variables were significantly correlated with TDASCM and TMINSP, 9 variables were significantly correlated with PAHINP and PSTONE, 8 variables were significantly correlated with PINTOL, 6 variables were significantly correlated with PTOPCA, and 2 variables were significantly correlated with PBLKAN at $\alpha = 0.05$. Nutrients, land use, road densities, and sedimentation appeared to be the most significant and strongly related variables to Ozark Highland IBI metrics (Table 2.9).

Dissolved oxygen, pH, water temperature (C), sulfates, ammonia, chlorides, nitrates, ortho-phosphates, total phosphorous, total organic carbon, turbidity, total

suspended solids, total dissolved solids, percent forested, percent urban, percent 100 m buffer forested, percent 100 m buffer urban, road density, and 100 m buffer road density contained at least 48 observations and were included in the PCA analysis. Principal component (PC) 1 accounted for 24.2% of model variance, while PCs 2, 3, 4 and 5 accounted for 23.3, 14.0, 9.9, and 9.8 % of model variance. Components 1, 2, 4 and 5 provided interpretable results. Percent forested, percent 100 m buffer forested, percent urban, percent 100 m buffer urban, and road density loaded heavily (loading > 0.800) on PC 1. Sulfates, chlorides, ortho-phosphates, and total phosphorous were heavily loaded on PC 2. Ammonia was heavily loaded on PC 3, and total suspended solids and turbidity (loading = 0.788) were heavily and moderately loaded on PC 4, respectively. Dissolved oxygen and temperature were heavily loaded on PC 5 (Table 2.10).

Discussion

We used existing and novel methods to develop an IBI for wadeable streams in Arkansas' Ozark Highlands ecoregion. In generating the IBI we used stringent metric-selection criteria to select 10 of 39 candidate metrics. Some of the metrics we selected are commonly used in IBI development (e.g., TSPECI and PINTOL), while others are unique to this study (e.g., PAHINP and PGBYCC). Our IBI metrics include taxonomic, functional, trophic, and reproductive fish-assemblage characters. Although we only selected 10 metrics, we believe the rigor of our metric-selection process selected IBI metrics with proven abilities to discern site quality, and our metric-contribution results indicate quality IBI metrics. For these reasons, and the fact that the data used represented

all three major Arkansas drainages in the ecoregion, we believe that the Ozark Highlands IBI is a robust index.

Comparing both the ADEQ and dual criteria reference and non-reference site classifications with IBI scoring classifications of reference sites revealed that 60.4 – 67.5% of all sites, including removed sites, were classified consistently. Only one site pre-classified as a reference site received an IBI score of less than 80. Therefore, all inconsistent classifications, except one, were from sites originally pre-classified as non-reference sites being classified as reference sites by the IBI. The one pre-classified reference site not classified as such by the IBI may be because it was an outlier that really did not reflect reference conditions. It may also have resulted from classifying only the upper or lower 50th percentile, as opposed to the 25th and 75th percentiles, of reference-site data for each metric as representing reference conditions. Even so, that site received a score of 79, which is within the scoring range between "reference" and "good" qualitative site classifications (i.e., 78 to 82) to be used for professional judgement of stream-site quality.

The inconsistent classifications may have been due to spatiotemporal variability in Ozark Highland streams, or that the site classifications were all or partially based on professional judgement. We lacked complete physiochemical data for all sites, and also lacked knowledge of truly pristine conditions in the ecoregion, to accurately identify reference sites. Also, our additional criterion of 75% forested watersheds may not have been the best criterion. Using additional variables such as riparian zone width, road density, in-stream habitat, etc. may have represented better reference-site selection criteria (Richards et al. 1996; Wang et al. 1997). Although, additional criteria would

have further reduced the total number of reference sites identified (i.e., 13), which would limit the ability of our reference sites to capture the natural variability of reference-site stream-fish-assemblage structure. Another possible reason for inconsistent site classifications is that only 23 sites contained data for the PBLKAN metric. This metric strongly influenced reference-site IBI scores (Table 2.8), and sites lacking this metric's data may not have reflected disturbances detected by this metric. Another reason for classification disagreement may be because most ADEQ fish collections represented good quality sites (W. Kieth, ADEQ, pers. com.), while only a few moderate to severely impacted sites were sampled (Figure 2.8). In fact, none of the relatively degraded sites that we sampled were classified as reference sites by the IBI, and some, as expected, scored poorly. These results indicate that the IBI can clearly differentiate good site quality from poor site quality. Nevertheless, the aforementioned factors may explain why the IBI classified many pre-classified non-reference sites as reference sites.

If there was error in selecting reference sites, is was in a conservative direction. It is preferable to leave out some sites scoring as reference sites from IBI development procedures than to include sites not scoring as reference sites. Including non-reference stream sites in development procedures would defeat the purpose of calibrating metrics to reference condition, which is an important component in IBI development. This conservative error, if it exists, treats the regulated community fairly. For example, it is better to misclassify a private landowner's stream as "unimpaired" rather than "impaired," because such a misclassification will not result in an undeserved regulatory burden of the landowner by regulatory agencies.

Within-site precision of the Ozark Highlands IBI was more precise than that of the IBI investigated by Fore et al. (1994), which was developed for Ohio's Great Miami basin. Our maximum 95% confidence-interval length was 3.2 scoring units, whereas Fore et al.'s (1994) was 12. In addition, the IBI scores used in Ohio's Great Miami basin range from 12 to 60, while Ozark Highland IBI scores range from 0 to 100. When Fore et al.'s (1994) IBI is standardized to score from 0 to 100, their maximum confidence-interval length is 25 units. Fore et al. (1994) also found that the precision of IBI scores was similar for least-disturbed sites using bootstrap data and sites sampled multiple times. This was not true for their disturbed sites, where variability in IBI scores increased for degraded sites using data from sites sampled multiple times, but variability remained constant across scores computed from bootstrap samples. Data were not available to assess annual variation in Ozark Highland IBI scores, but Hughes et al. (1998) found their IBI was able to detect an 8% (8 unit) change in IBI scores between years. Fore et al. (1994) found their IBI was able to detect an 18% (8.5 units) difference in IBI scores between years. Karr et al. (1987) found their IBI scores were concordant in ranking sites over time.

Our bootstrap sample data indicated no relation between site quality (i.e., mean IBI score) and the confidence-interval lengths of those sites. Fore et al. (1994) reported similar findings for their bootstrap data. Conversely, they found least-disturbed site IBI scores varied least across years, a result also reported by Karr et al. (1987). Hughes et al. (1998) reported a weak increase in IBI score standard deviations with increasing IBI scores, although variance was greatest at intermediate IBI scores. There appears to be a negative, but insignificant and variable, trend between mean IBI scores and confidence-

interval lengths of those scores for our bootstrap data (Figure 2.9). We also found no relation between IBI-score confidence-interval length and number of individuals and species richness, results that contradict findings reported by Fore et al. (1994).

Metric contribution analyses indicated that quality metrics were chosen. Our variances of differences between IBI scores with and without individual metrics averaged 6.60 (SD = 3.31), whereas metrics selected by Bowen et al. (1998) and Minns et al. (1994) averaged 5.84 (SD = 4.23) and 6.30 (SD = 2.04) respectively. This may indicate that our metrics' individual contributions to the IBI are greater than those used by Bowen et al. (1998) and Minns et al. (1994). Although, Bowen et al.'s (1998) IBI included nine metrics, and Minns et al.'s (1994) IBI included 12 metrics. Unequal numbers of metrics between IBIs may explain why our average metric contribution using the variance method was higher than Minns et al.'s (1994) metrics. Results from the metric contribution analysis indicated that trophic metrics contributed the most to IBI scores, a result also found by Bowen et al. (1998) and Minns et al. (1994). Reproductive metrics were the least influential on IBI scores. Although this result is confounded somewhat because the taxonomic metric PGBYCC, which represented tolerant species in the region, also contains three out of four species in the reproductive classification 7 (miscellaneous, site-prep, parental-care spawners). PGBYCC ranked as the 6th most influential metric in the IBI. Therefore, reproductive factors may have contributed to the IBI to a greater extent than is indicated by TMINSP itself.

Using data from sampled reaches greater or less than the 51 MSWs used to develop the IBI can significantly affect IBI scores. Although, IBI scores at 50 MSWs (the closest number of MSWs to 51) were not significantly different from scores

calculated for sampling distances of 40 to 55 MSWs. Therefore, there may be some acceptable deviation from sampling 51 MSWs. In addition, the within-site precision range of 4 units is, on average, larger than most average IBI score deviations for different sampling lengths when compared to IBI scores at 50 MSWs (Figure 2.7). Although, sampling less than 40 MSWs or greater than 55 MSWs may result in IBI scores differing significantly from IBI scores calculated at 51 MSWs, and may adversely and inconsistently affect IBI scores. Even a 5 MSW deviation from 51 MSWs may potentially cause qualitative site classifications to change approximately 10% of the time (Figure 2.7). In addition, all taxonomic richness metrics were affected by sampling effort, as expected, since taxonomic richness can only increase with sampling effort. Therefore, future IBIs may benefit from excluding certain metrics, such as taxonomic richness metrics that are strongly affected by sampling effort, especially if they are not major contributors to IBI scores (see Table 2.8). More research is needed to weigh metric contributions to IBI scores against the probability that particular metrics add an unacceptable level of uncertainty to the IBI. Others (e.g., Angermeier and Karr (1986)) have also shown sampling effort to affect IBI scores. Further research on the effects of sampling effort on our IBI scores is warranted, especially regarding the use of various gear types (e.g., barge electrofishing). Although each lotic system is unique and fish assemblage characters such as species accumulation rates vary among systems, a standardizing sampling effort of 51 MSW when sampling using backpack electrofishing, and an equivalent effort using barge electrofishing, will result in the most consistent and accurate IBI scores. Therefore, we recommend sampling 51 MSW, or an equivalent

sampling effort in larger streams, using a one-pass electrofishing sample when applying this IBI to wadeable, Ozark Highland streams.

Our sampling effort designation for the Ozark Highlands IBI is greater than what is recommended for most bioassessment purposes (see Barbour et al. (1999) and Meador et al. (1993)). We think that this sampling distance is important for an accurate characterization of stream-fish assemblages and stream health. Although a relatively short distance is needed to estimate assemblage structure, sampling a distance less than the designated effort will, on average, result in additional species not being sampled. Missing multiple species may result in taxonomic richness metrics (e.g., TSPECI) failing to accurately, and consistently, reflect stream health. Because it is not known which species will be missed during a sampling episode (Chapter 1), missing different species during each sampling episode may contribute variability to taxonomic richness metrics, and thus IBI scores. Also, proportional metrics representing a low proportion of the fish assemblage may be affected if the species comprising that metric are not sampled. For example, in our IBI the PTOPCA metric is composed of top carnivores, which has an upper threshold of approximately three percent. Because top carnivores are usually larger individuals, they generally occupy deep pool habitats. All pools in Ozark Highland streams do not appear to contain suitable habitat for top carnivores. Therefore, sampling a sufficient distance will increase the probability that the species comprising the PTOPCA metric will be collected, if any suitable top-carnivore habitat is present at the sample site. For these reasons, accurate fish assemblage characterization and sampling consistency is crucial for determining stream health using an IBI (Fausch et al. 1990).

All metrics were shown to be correlated with at least two of the selected physiochemical and land-use variables. Nutrients, land use, road density, and sedimentation levels were most consistently correlated with IBI metrics. Similar relations have been shown in other U.S. regions and are thought to be the most important factors contributing to stream degradation in the U.S. (Matthews et al. 1992; Lenat and Crawford 1994; Wang et al. 1997; Lammert and Allan 1999; Waite and Carpenter 2000). Also, land-use proximity to streams is not as strongly related as entire watershed land use to the IBI metrics. Mechanisms influencing these relations may be similar to those discussed by Omernik et al. (1981). They claim that land use adjacent to streams may act as a nutrient sink for a period of time, but eventually nutrient inflow must equal nutrient outflow. In addition, unmapped waterways and storm runoff may contribute nutrients and sediments from areas away from lands bordering streams.

The PCA revealed some expected relations among physiochemical and land-use variables. For example, PC 1, which explained 24.2% of the variance, indicated a positive relation between road density and percent urban land-use and a negative relation between percent forested land-use and percent urban land-use and road density. In addition, nutrient concentrations showed a moderate positive relation with PC 1 (Table 2.10). This may result from all agricultural land-use identified in the Ozark Highland watersheds as being pastureland. Nutrients from fertilizers or livestock excretory products may be absorbed or buffered by pasture vegetation. Also, nutrients may make their way into streams predominantly during episodic events such as storms producing runoff. Because we sampled during base-flow conditions, we may not have detected the true nutrient load resulting from agricultural land use. Both phosphorous variables, as

well as sulfates and chlorides, were heavily loaded on PC 2. These variables should be positively related, if a relation exists, and may occur from high nutrient inputs, organic decay, and possibly water-treatment effluents. The Arkansas Department of Pollution Control & Ecology (ADPCE 1993) reported coincident high concentrations of phosphorous, chlorides, and TDS from an untreated wastewater-treatment-facility discharge, and higher chloride and sulfate levels below, as opposed to above, treated effluent discharges. Although data from that particular untreated discharge event were not used in analyses, it demonstrates that untreated effluents may be causing the relations observed between phosphorous, chlorides, and sulfates. Observed relations between ammonia, pH, and total dissolved solids may be signals of natural watershed conditions, but otherwise remain unexplained. Relations between turbidity and total suspended solids on PC 4, and between temperature and dissolved oxygen on PC 5, conform to water-quality expectations (Allan 1995).

Our development of the Ozark Highlands IBI seemed to provide a beneficial use of an existing database. Although the database was supplemented, using existing data had advantages and disadvantages for IBI development. Some of the data already existed in a form amenable to IBI development. However, it may not have been collected with the consistency that is most desirable for IBI development (Karr et al. 1986). The length of stream the ADEQ sampled is probably not as consistent as the sites sampled to compliment the existing fish collections because one of our objectives for additional sampling was to sample in a standardized manner. The ADEQ strives for consistency in that they sample very intensively in an attempt to sample in all habitat types and collect all species present. Although this approach meets their objectives, and is desirable for

IBI development (Angermeier and Karr 1986), it may have led to inconsistencies in sampling effort among fish collections across the region. We worked with ADEQ field crews to gain insight into their sampling effort. Based on our observations of ADEQ field crews and field measurements of stream distance based on MSW, we believe that our sampling scheme closely approximated the ADEQ field crew's sampling procedures. Thus, our recommended sampling effort should consistently characterize Ozark Highland stream-fish assemblages, and provide a scheme that consistently reflects stream health.

Index of Biotic Integrity development for the Ozark Highlands ecoregion has provided a valuable use of existing data that has resulted in a useful tool for resource managers. The caveats of this study are comparable to other studies using data not collected for a single purpose. Although most of the data we used were not intended for this study, we think these data, because of thorough sampling regimes, provided an excellent opportunity to develop an IBI that can differentiate stream-site conditions in Arkansas' Ozark Highlands. Therefore, our IBI can be used to determine stream-site quality, identify impaired streams, and provide a means to prioritize stream-rehabilitation efforts and meet CWA requirements.

Acknowledgements

This manuscript was funded through a sub-contract agreement with Tetra Tech, Inc. via the U.S. Environmental Protection Agency contract #68–C98–111 Work Assignment 1-10. We thank all ADEQ personnel involved in their fish collections and data management. We are grateful to W. Keith, J. Wise, B. Posey, and C. Davidson for

classifying stream condition based on their professional judgements. W. Keith made important contributions to this work through his comments and insights on past fish collections and their associated data. R. Smogor provided insights and ideas on IBI development and the effects of sampling effort. P. Crocker, M. Barbour, N. Stone and E. Buckner commented on earlier drafts of this chapter. We also thank C. Okiror, J. Poole, T. Hungerford, L. Lackey, and S. Mutagyera for help with data collection and entry.

References

ADPCE. 1993. Water quality survey of the Huntsville STP effluent on Town Branch and Holman Creek. Arkansas Department of Pollution Control & Ecology, Little Rock, Arkansas. WQ93-03-1.

ADPCE. 1995. Water quality, macroinvertebrate, and fish community survey of the Upper White River watershed, northwest Arkansas. Arkansas Department of Pollution Control & Ecology, Little Rock, Arkansas. WQ95-08-1.

ADPCE. 1996. Arkansas water quality inventory report. Arkansas Department of Pollution Control & Ecology, Little Rock, Arkansas.

Allan, J. D. 1995. Stream ecology: structure and function of running waters. Chapman & Hall, New York, New York.

Angermeier, P. L. and J. R. Karr. 1986. Applying an index of biotic integrity based on stream-fish communities: considerations in sampling and interpretation. North American Journal of Fisheries Management 6:418-429.

Angermeier, P. L., R. A. Smogor, and J. R. Stauffer. 2000. Regional frameworks and candidate metrics for assessing biotic integrity in Mid-Atlantic highland streams. Transactions of the American Fisheries Society 129:962-981.

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in wadeable streams and rivers: periphyton, benthic macroinvertebrates, and fish, Second edition. U.S. Environmental Protection Agency; Office of Water; Washington D.C. EPA 841-B-99-002.

Barbour, M. T., J. L. Plafkin, B. P. Bradley, C. G. Graves, and R. W. Wisseman. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: metric redundancy and variability among reference sites. Environmental Toxicology and Chemistry 11:437-449.

Bowen, Z. H., M. C. Freeman, and D. L. Watson. 1998. Index of Biotic Integrity applied to a flow-regulated river system. Proceedings of the Annual Conference of the Southeast Association of Fish and Wildlife Agencies 50:26-37.

Didier, J. and P. Kestemont. 1996. Relationships between mesohabitats, ichthyological communities, and IBI metrics adapted to a European river basin (The Meuse, Belgium). Hydrobiologia 341:133-144.

Etnier, D. A. and W. C. Starnes. 1993. The fishes of Tennessee, 1st ed. The University of Tennessee Press, Knoxville, Tennessee.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional application of an Index of Biotic Integrity based on stream fish communities. Transactions of the American Fisheries Society 113:39-55.

Fausch, K. D., J. Lyons, J. R. Karr, and P. L. Angermeier. 1990. Fish communities as indicators of environmental degradation. Pages 123-144 *in* S. M. Adams, editor. Biological indicators of stress in fish. American Fisheries Society Symposium 8, Bethesda, Maryland.

Fore, L. S., J. R. Karr, and L. L. Conquest. 1994. Statistical properties of an Index of Biological Integrity used to evaluate water resources. Canadian Journal of Fisheries and Aquatic Sciences 51:1077-1087.

Gauch, H. G., Jr. 1982. Multivariate analysis in community ecology. Cambridge University Press, New York.

Giese, J., B. Keith, M. Maner, R. McDaniel, and B. Singleton. 1987. Physical, chemical, and biological characteristics of least-disturbed reference streams in Arkansas' ecoregions: data analysis. Arkansas Department of Pollution Control & Ecology, Little Rock.

Hlass, L. J., W. L. Fisher, and D. J. Turton. 1998. Use of the Index of Biotic Integrity to assess water quality in forested streams in the Ouachita Mountains ecoregion, Arkansas. Journal of Freshwater Ecology 13:181-192.

Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences 55:1618-1631.

Hughes, R. M., E. Rexstad, and C. E. Bond. 1987. The relationship of aquatic ecoregions, river basins and physiographic provinces to the ichthyogeographic regions of Oregon. Copeia 1987:423-432.

Jenkins, R. E. and N. M. Burkhead. 1993. Freshwater fishes of Virginia. American Fisheries Society, Bethesda, Maryland.

Jones, J. R., B. H. Tracy, J. L. Sebaugh, D. H. Hazelwood, and M. M. Smart. 1981. Biotic index tested for ability to assess water quality of Missouri Ozark streams. Transactions of the American Fisheries Society 110:627-637.

Karr, J. R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running water: a method and its rationale. Illinois Natural History Survey, Champaign, Illinois. Special Publication 5.

Karr, J. R., P. R. Yant, K. D. Fausch, and I. J. Schlosser. 1987. Spatial and temporal variability of the Index of Biotic Integrity in three midwestern streams. Transactions of the American Fisheries Society 116:1-11.

Keith, W. E. 1987. Distribution of fishes in reference streams within Arkansas' ecoregions. Proceedings Arkansas Academy of Science 41:57-60.

Lammert, M. and J. D. Allan. 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23:257-270.

Lenat, D. R. and J. K. Crawford. 1994. Effects of land use on water quality and aquatic biota of three North Carolina Piedmont streams. Hydrobiologia 294:185-199.

Lyons, J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. North American Journal of Fisheries Management 12:198-203.

Lyons, J., S. Navarro-Perez, P. A. Cochran, E. Santana, and M. Guzman-Arroyo. 1995. Index of Biotic Integrity based on fish assemblages for the conservation of streams and rivers in west-central Mexico. Conservation Biology 9:569-584.

Lyons, J., L. Wang, and T. D. Simonson. 1996. Development and validation of an Index of Biotic Integrity for coldwater streams in Wisconsin. North American Journal of Fisheries Management 16:241-256.

Matthews, W. J., D. J. Hough, and H. W. Robison. 1992. Similarities in fish distribution and water quality patterns in streams of Arkansas: congruence of multivariate analyses. Copeia 1992:296-305.

Matthews, W. J. and H. W. Robison. 1988. The distribution of the fishes of Arkansas: a multivariate analysis. Copeia 1988:358-374.

Meador, M. R., T. F. Cuffney, and M. E. Gurtz. 1993. Methods for sampling fish communities as part of the National Water Quality Assessment. U.S. Geological Survey Open File Report, Raleigh, North Carolina. USGS/OPR 93-104.

Miller, D. L., P. M. Leonard, R. M. Hughes, J. R. Karr, P. B. Moyle, L. H. Schrader, B. A. Thompson, R. A. Daniels, K. D. Fausch, G. H. Fitzhugh, J. R. Gammon, D. B. Halliwell, P. L. Angermeier, and D. J. Orth. 1988. Regional applications of an Index of Biotic Integrity for use in water resource management. Fisheries 13:12-20.

Minns, C. K., V. W. Cairns, R. G. Randall, and J. E. Moore. 1994. An Index of Biotic Integrity (IBI) for fish assemblages in the littoral zone of Great Lakes' areas of concern. Canadian Journal of Fisheries and Aquatic Sciences 51:1804-1822.

Ohio EPA. 1987. Biological criteria for the protection of aquatic life: volume II: users manual for biological field assessment of Ohio surface waters. State of Ohio Environmental Protection Agency, Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Omernik, J. M., A. R. Abernathy, and L. M. Mate. 1981. Stream nutrient levels and proximity of agricultural and forest land to streams: some relationships. Journal of Soil and Water Conservation 36:227-231.

Paller, M. H. 1995. Relationships among number of fish species sampled, reach length surveyed, and sampling effort in South Carolina Coastal Plain streams. North American Journal of Fisheries Management 15:110-120.

Patton, T. M., W. A. Hubert, F. J. Rahel, and K. G. Gerow. 2000. Effort needed to estimate species richness in small streams in the Great Plains in Wyoming. North American Journal of Fisheries Management 20:394-398.

Pflieger, W. L. 1997. The fishes of Missouri, revised ed. Missouri Department of Conservation, Jefferson City, Missouri.

Richards, C. L., L. B. Johnson, and G. E. Host. 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53:295-311.

Robison, H. W. and T. M. Buchanan. 1988. Fishes of Arkansas. The University of Arkansas Press, Fayetteville, Arkansas.

Rohm, C. M., J. W. Giese, and C. C. Bennett. 1987. Evaluation of an aquatic ecoregion classification of streams in Arkansas. Journal of Freshwater Ecology 4:127-141.

Simon, T. P. and J. Lyons. 1995. Application of the Index of Biotic Integrity to evaluate water resource integrity in freshwater ecosystems. Pages 245-262 *in* W. S. Davis and T. P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. CRC Press, Inc., Boca Raton, Florida.

Smith, K. G., R. S. Dzur, D. G. Catanzaro, M. E. Garner, and W. F. Limp. 1998. The Arkansas GAP analysis project: final report. Center for Advanced Spatial Technologies (CAST), Fayetteville, Arkansas.

Smogor, R. A. 1996. Developing an Index of Biotic Integrity (IBI) for warmwater wadeable streams in Virginia. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Smogor, R. A. and P. L. Angermeier. 1999. Effects of drainage basin size and anthropogenic disturbance on relations between stream size and IBI metrics in Virginia. Pages 249-272 *in* T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Inc., Boca Raton, Florida.

Sokal, R. R. and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research, 3rd ed. W. H. Freeman and Company, New York.

Waite, I. R. and K. D. Carpenter. 2000. Associations among fish assemblage structure and environmental variables in Willamette Basin Streams, Oregon. Transactions of the American Fisheries Society 129:754-770.

Wang, L., J. Lyons, and R. Gatti. 1997. Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries 22:6-12.

Ward, R. C., J. C. Loftis, and G. B. McBride. 1986. The "data-rich but information-poor" syndrome in water quality monitoring. Environmental Management 10:291-297.

Weih, R. 2001. Natural state digital database: version 1.0 Beta 5, UTM Zone 15, NAD27. Arkansas Forest Resource Center, University of Arkansas-Monticello, Monticello, Arkansas.

Zar, J. H. 1999. Biostatistical analysis, 4th ed. Prentice Hall, Upper Saddle River, New Jersey.

Table 2.1. Fish classifications used to derive metric values. Number designations are described in the Candidate IBI Metrics section.

Family/Species	Top Carnivore	Intolerant	Benthic	Trophic	Repro-ductive
Petromyzontidae					
Ammocoete				1	2
Lepisosteidae					
Lepisosteus osseus	x			8	8
Clupeidae					
Dorosoma cepedianum				2	8
Cyprinidae					
Campostoma anomalum			x	4	2
Campostoma oligolepis			x	1	2
Ctenopharyngodon idella				4	8
Cyprinella galactura		x		6	8
Cyprinella whipplei		x		6	8
Cyprinus carpio				2	8
Erimystax dissimilis		x	x	2	1
Erimystax harryi			x	2	1
Erimystax x-punctatus		x	x	2	1
Hybopsis amblops		x	x	6	1
Luxilus cardinalis		x		4	2
Luxilus chrysocephalus				6	2
Luxilus pilsbryi		x		4	2
Luxilus zonatus				6	2
Lythrurus umbratilis				4	2
Nocomis asper		x		4	2
Nocomis bigutattus		x		4	2
Notemigonus crysoleucas				4	8
Notropis boops		x		6	1
Notropis greenei		x		6	1
Notropis nubilus				1	2
Notropis ozarcanus				6	1
Notropis rubellus		x		4	1
Notropis telescopus		x		6	1
Phoxinus erythrogaster		x		1	2
Pimephales notatus				4	7
Pimephales promelas				2	6
Pimephales tenellus		x		4	3
Semotilus atromaculatus				7	2
Catostomidae					
Carpiodes cyprinus			x	2	8
Carpiodes velifer			x	2	8
Catostomus commersoni		x	x	6	1

Table 2.1. Continued.

Family/Species	Top Carnivore	Intolerant	Benthic	Trophic	Repro-ductive
Erimyzon oblongus			x	4	2
Hypentelium nigricans		x	x	4	1
Ictiobus bubalus				2	8
Ictiobus niger				4	8
Minytrema melanops			x	2	1
Moxostoma carinatum		x	x	6	2
Moxostoma duquesnei		x	x	6	1
Moxostoma erythrurum			x	4	1
Moxostoma macrolepidotum			x	6	1
Ictaluridae					
Ameiurus melas				3	4
Ameiurus natalis				5	7
Ictalurus punctatus				5	7
Noturus albater		x	x	6	4
Noturus exilis		x	x	6	4
Noturus flavater		x	x	7	4
Pylodictis olivaris	x			8	7
Salmonidae					
Oncorhyncus mykiss		x		7	2
Salmo trutta		x		7	2
Aphredoderidae					
Aphredoderus sayanus				6	8
Fundulidae					
Fundulus catenatus		x		6	3
Fundulus notatus				4	6
Fundulus olivaceus				6	8
Poecilliidae					
Gambusia affinis				4	8
Atherinidae					
Labidesthes sicculus				6	8
Cottidae					
Cottus carolinae		x	x	7	4
Cottus hypselurus		x	x	7	4
Percichthyidae					
Morone chrysops	x			8	8
Centrarchidae					
Ambloplites ariommus	x	x		7	4
Ambloplites constellatus	x	x		7	4
Ambloplites rupestris	x	x		7	4
Lepomis cyanellus				7	7
Lepomis gulosus				7	7

Table 2.1. Continued.

Family/Species	Top Carnivore	Intolerant	Benthic	Trophic	Repro-ductive
Lepomis macrochirus				5	7
Lepomis megalotis				6	2
Lepomis microlophus				6	7
Lepomis punctatus				4	7
Micropterus dolomieu	x	x		7	4
Micropterus punctulatus	x			7	7
Micropterus salmoides	x			7	7
Pomoxis annularis	x			7	7
Pomoxis nigromaculatus	x			7	7
Percidae					
Etheostoma blennioides		x	x	6	8
Etheostoma caeruleum		x	x	6	1
Etheostoma euzonum		x	x	6	1
Etheostoma flabellare		x	x	6	4
Etheostoma juliae		x	x	6	1
Etheostoma punctulatum		x	x	6	1
Etheostoma spectabile			x	6	1
Etheostoma stigmaeum		x	x	6	1
Etheostoma zonale		x	x	6	8
Percina caprodes			x	6	1
Percina evides		x	x	6	1
Percina nasuta		x	x	6	1
Stizostedion vitreum	x	x		7	1

Table 2.2. Metric acronyms and definitions. Percent metrics represent the percent of individuals sampled. Precise definitions are included in the Candidate IBI Metrics section.

Acronym	Definition
INDSEC	Number of individuals per second electrofishing
PAHINP	Percent as algivores/herbivores, invertivores, and piscivores
PAHINV	Percent as algivores/herbivores and invertivores
PANOMA	Percent with an anomaly
PBENTH	Percent as benthics
PBLK3+	Percent with 3+ black spot cysts
PBLKAN	Percent with a black spot cyst or an anomaly
PBLKSP	Percent with a black spot cyst
PCYPRI	Percent as cyprinids
PDARSC	Percent as darters and sculpins
PDASCM	Percent as darters, sculpins, and madtoms
PDETAH	Percent as detritivores and algivores/herbivores
PGBYCC	Percent as green sunfish, bluegill, yellow bullhead, and channel catfish
PGENER	Percent as generalists
PGRESF	Percent as green sunfish
PGSFYB	Percent as green sunfish and yellow bullhead
PINTOL	Percent as intolerants
PINVER	Percent as invertivores
PINVPI	Percent as invertivores and piscivores
PMINSP	Percent as mineral, site-prep spawners
PSHINE	Percent as shiners
PSIMLI	Percent as simple, lithophilic spawners
PSIMPX	Percent as simple, miscellaneous spawners
PSTONE	Percent as stonerollers
PTOPCA	Percent as top carnivores
PXSPPC	Percent as miscellaneous, site-prep, parental-care spawners
TBENTH	Total number of benthic species
TCYPRI	Total number of cyprinid species
TDARSC	Total number of darter and sculpin species
TDARTE	Total number of darter species
TDASCM	Total number of darter, sculpin, and madtom species
TGENER	Total number of generalist species
TINTOL	Total number of intolerant species
TMINSP	Total number of mineral spawning species
TSHINE	Total number of shiner species
TSIMLI	Total number of simple, lithophilic spawning species
TSPECI	Total number of species
TSUCKE	Total number of sucker species
TSUNFI	Total number of sunfish species

Table 2.3. Results of t-tests between metric values of reference and non-reference sites. All t-tests were computed assuming equal or unequal variances depending on F-test results. Metrics were considered significantly different at $\alpha = 0.10$. Significant metrics chosen for additional analyses as candidate metrics are indicated. For all metrics reference n = 10 and non-reference n = 67, except PBLKSP and PBLKAN where reference n = 6 and non-reference n = 13.

	t-test			t-test	
Metric	p-value	Selected	Metric	p-value	Selected
INDSEC	0.617		PSHINE	0.290	
PAHINP	0.054	x	PSIMLI	0.588	
PAHINV	0.301		PSIMPX	0.131	
PANOMA	0.574		PSTONE	0.054	x
PBENTH	0.324		PTOPCA	0.010	x
PBLK3+	0.256		PXSPPC	0.002	x
PBLKAN	0.005	x	TBENTH	0.058	x
PBLKSP	0.058	x	TCYPRI	0.336	
PCYPRI	0.171		TDARSC	0.045	x
PDARSC	0.708		TDARTE	0.053	x
PDASCM	0.449		TDASCM	0.043	x
PDETAH	0.575		TGENER	0.783	
PGBYCC	0.011	x	TINTOL	0.010	x
PGENER	0.131		TMINSP	0.023	x
PGRESF	0.016	x	TSHINE	0.138	
PGSFYB	0.040	x	TSIMLI	0.372	
PINTOL	0.035	x	TSPECI	0.082	x
PINVER	0.051	x	TSUCKE	0.227	
PINVPI	0.724		TSUNFI	0.806	
PMINSP	0.711				

Table 2.4. Pearson's correlations (r) between candidate metrics discriminating between reference and non-reference sites. Redundant metrics are indicated by an r < -0.90 or r > 0.90. All metrics n = 77, except PBLKSP and PBLKAN where n = 19.

Metric	PBLKAN	PBLKSP	PGBYCC	PGRESF	PGSFYB	PINTOL	PINVER	PSTONE	PTOPCA
PAHINP	-0.029	-0.053	0.697	0.392	0.439	-0.362	-0.123	-0.056	-0.158
PBLKAN		0.979	0.433	0.634	0.611	-0.463	0.062	0.156	-0.174
PBLKSP			0.370	0.564	0.538	-0.395	0.028	0.111	-0.214
PGBYCC				0.933	0.945	-0.496	-0.117	-0.004	-0.226
PGRESF					0.991	-0.452	-0.088	0.023	-0.211
PGSFYB						-0.486	-0.086	0.039	-0.218
PINTOL							0.123	-0.675	0.303
PINVER								-0.464	0.182
PSTONE									-0.251

Metric	PXSPPC	TBENTH	TDARSC	TDARTE	TDASCM	TINTOL	TMINSP	TSPECI
PAHINP	0.729	-0.246	-0.273	-0.227	-0.283	-0.289	-0.267	-0.080
PBLKAN	0.368	-0.298	-0.231	-0.253	-0.259	-0.290	-0.365	-0.379
PBLKSP	0.313	-0.207	-0.122	-0.181	-0.184	-0.262	-0.340	-0.375
PGBYCC	0.945	-0.432	-0.424	-0.356	-0.448	-0.473	-0.475	-0.308
PGRESF	0.845	-0.431	-0.407	-0.343	-0.433	-0.461	-0.474	-0.356
PGSFYB	0.864	-0.449	-0.422	-0.356	-0.450	-0.474	-0.485	-0.356
PINTOL	-0.508	0.409	0.450	0.345	0.477	0.447	0.409	0.234
PINVER	-0.021	0.245	0.268	0.409	0.338	0.366	0.347	0.328
PSTONE	-0.061	-0.250	-0.251	-0.227	-0.288	-0.307	-0.295	-0.232
PTOPCA	-0.220	0.412	0.419	0.410	0.488	0.524	0.483	0.415
PXSPPC		-0.372	-0.382	-0.294	-0.393	-0.398	-0.408	-0.212
TBENTH			0.945	0.920	0.935	0.922	0.923	0.900
TDARSC				0.962	0.975	0.870	0.853	0.809
TDARTE					0.955	0.861	0.843	0.837
TDASCM						0.901	0.865	0.817
TINTOL							0.901	0.817
TMINSP								0.927

Table 2.5. Linear regression results of watershed size (independent) versus IBI metrics (dependent). Regressions were considered significant at $\alpha = 0.05$. Metrics significantly affected by watershed size require linear functions to determine threshold limits for metric scoring.

Metric	n	r^2	p-value
PAHINP	10	0.033	0.614
PBLKAN	6	0.039	0.708
PGBYCC	10	0.030	0.634
PINTOL	10	0.114	0.341
PINVER	10	0.015	0.738
PSTONE	10	0.018	0.710
PTOPCA	10	0.151	0.267
TDASCM	10	0.622	0.007
TMINSP	10	0.559	0.013
TSPECI	10	0.512	0.020

Table 2.6. IBI metrics, their intercept (A) and slope (B) scoring coefficients, and their upper (UT) and lower (LT) threshold limits. Metrics requiring a scoring adjustment for watershed size (km^2; WS) have linear functions to derive upper threshold limits. Metric scores (MS) are calculated from raw metric values (MR) using the equation: MS = A + B*(MR). Raw metric values above or below threshold limits take the value of their respective threshold limit. See Metric Scoring section for scoring instructions.

Metric	Metric Coefficients		Threshold Limits	
	A	B	LT	UT
PAHINP	10.24	-10/9.00	0.22	9.22
PBLKAN	11.53	-10/0.57	0.09	0.66
PGBYCC	10.15	-10/24.99	0.38	25.37
PINTOL	0	10/57.10	0	57.10
PINVER	0	10/40.00	0	40.00
PSTONE	18.34	-10/38.43	32.10	70.53
PTOPCA	0	10/2.80	0	2.80
TDASCM[a]	0	10/12.00	0	12.00
TMINSP[a]	0	10/26.00	0	26.00
TSPECI[a]	0	10/35.00	0	35.00
TDASCM[b]	0	10/UT - LT	0	4.909 + 0.0091 (WS)[c]
TMINSP[b]	0	10/UT - LT	0	14.336 + 0.0150 (WS)[c]
TSPECI[b]	0	10/UT - LT	0	19.218 + 0.0199 (WS)[c]

[a]Metric coefficients and threshold values for sites with watershed sizes (WS) ≥ 800 km^2
[b]Metric coefficients and threshold values for sites with watershed sizes (WS) < 800 km^2
[c]Round upper threshold to nearest integer.

Table 2.7. Percent consistency values for reference and non-reference site classifications. "IBI reference" indicates sites with an IBI score from 80 – 100. "Pre-classified" indicates the author's (n = 96) and the ADEQ's (n = 80) site classifications prior to IBI development. Removed sites represent 19 randomly chosen sites not included in IBI development; three of the 19 sites were not classified by the ADEQ (n = 16). All sites included those used to develop the IBI and removed sites. The authors classified sites by using a combination of objective and subjective criteria (i.e., dual criteria). The ADEQ classifications were based on professional judgement.

Outcomes	Removed Sites		All Sites	
	Dual Criteria	ADEQ	Dual Criteria	ADEQ
Consistent Classifications	47.4	43.8	60.4	67.5
IBI reference, pre-classified as non-reference	47.4	50.0	38.5	28.8
IBI non-reference, pre-classified as reference	5.2	6.2	1.1	3.7

Table 2.8. Correlation coefficients (r) of simple linear correlations between metric scores and IBI scores computed without each metric, and variances of differences between IBI scores with and without each metric. Metric correlation coefficients and variances were ranked according to each metrics contribution to IBI scores. Values in the Rank column represent the ranks of the sums of the ranks of correlation coefficients and variances. Ranks indicate the relative contributions of metrics to the IBI. A rank of "1" indicates the highest relative contribution of a metric to the IBI. A rank of "10" indicates the lowest relative contribution of a metric to the IBI. Metric contributions were determined using all sites (n = 96) and sites with IBI scores in reference (n = 49), good (n = 26), fair (n = 15), and poor (n = 6) scoring ranges.

Classification	Metric	r	Variance	Rank
All Sites	PAHINP	0.300	8.57	1.5
	PBLKAN	0.403	6.98	4.5
	PGBYCC	0.511	5.13	6
	PINTOL	0.620	6.61	7
	PINVER	0.374	9.75	1.5
	PSTONE	0.480	7.21	4.5
	PTOPCA	0.516	12.93	3
	TDASCM	0.699	3.49	8.5
	TMINSP	0.798	1.96	10
	TSPECI	0.622	3.33	8.5
Reference	PAHINP	-0.230	3.50	4
	PBLKAN	-0.479	5.40	1.5
	PGBYCC	0.195	0.96	9
	PINTOL	0.134	3.18	6
	PINVER	-0.229	6.53	1.5
	PSTONE	-0.030	1.40	5
	PTOPCA	-0.016	9.49	3
	TDASCM	0.201	2.47	8
	TMINSP	0.251	1.30	10
	TSPECI	0.013	1.54	7
Good	PAHINP	-0.221	7.64	5
	PBLKAN	0.083	7.32	6.5
	PGBYCC	-0.138	4.88	6.5
	PINTOL	-0.310	11.03	2
	PINVER	-0.318	11.02	2
	PSTONE	-0.436	7.19	4
	PTOPCA	-0.304	16.35	2
	TDASCM	-0.157	4.31	8
	TMINSP	0.314	2.93	10
	TSPECI	0.016	4.57	9
Fair	PAHINP	-0.537	18.51	3
	PBLKAN	0.834	5.42	9.5

Table 2.8. Continued.

Classification	Metric	r	Variance	Rank
Fair	PGBYCC	-0.106	12.37	5
	PINTOL	-0.170	11.16	5
	PINVER	-0.638	19.58	2
	PSTONE	-0.654	20.85	1
	PTOPCA	-0.282	8.02	5
	TDASCM	0.322	6.03	7.5
	TMINSP	0.590	2.28	9.5
	TSPECI	0.272	5.51	7.5
Poor	PAHINP	-0.846	27.13	1
	PBLKAN	NA	NA	NA
	PGBYCC	-0.769	24.75	3
	PINTOL	0.137	2.24	8
	PINVER	-0.599	10.36	4
	PSTONE	-0.754	27.15	2
	PTOPCA	-0.447	6.89	5
	TDASCM	0.168	3.10	7
	TMINSP	0.417	2.89	9
	TSPECI	0.038	6.60	6

Table 2.9. Significant ($\alpha = 0.05$) Spearman's rank correlations (r_s) between IBI metrics and selected physiochemical and land-use variables. Sample sizes (n) ranged from 5 to 96 for each variable, and were 96 unless otherwise noted (in parentheses).

Metric	Variable	r_s	Metric	Variable	r_s
PAHINP	Zinc (18)	0.470	PSTONE	Chromium (12)	0.624
	Fluoride (16)	0.502		Copper (17)	0.576
	% Forested	-0.348		Chloride (55)	0.378
	% Agriculture	0.352		Nitrates (57)	0.315
	% Urban	0.282		O-Phosphates (54)	0.289
	% Buffer Forested	-0.316		TDS (55)	0.378
	% Buffer Agriculture	0.306		% Buffer Urban	0.206
	% Buffer Urban	0.231		Road Density	0.243
	Road Density	0.268		Sedimentation (26)	0.499
PBLKAN	% Forested (23)	-0.415	PTOPCA	Chloride (55)	-0.378
	Sedimentation (21)	0.579		Nitrates (57)	-0.456
				Turbidity (48)	0.319
PGBYCC	Sodium (17)	0.543		TDS (55)	-0.416
	Zinc (18)	0.639		Road Density	-0.292
	Fluoride (16)	0.556		Buffer Road Density	-0.281
	Chloride (55)	0.382			
	Nitrates (57)	0.332	TDASCM	Ammonia (54)	0.323
	O-Phosphates (54)	0.320		Chloride (55)	-0.581
	% Forested	-0.395		Nitrates (57)	-0.566
	% Agriculture	0.327		T-Phosphorous (57)	-0.312
	% Urban	0.253		Turbidity (48)	0.491
	% Buffer Forested	-0.297		TSS (55)	0.287
	% Buffer Agriculture	0.263		TDS (55)	-0.570
	Road Density	0.408		% Forested	0.310
	Buffer Road Density	0.256		% Agriculture	-0.211
	Sedimentation (26)	0.620		% Buffer Forested	0.213
				Road Density	-0.413
PINTOL	Calcium (17)	0.502		Buffer Road Density	-0.283
	Potassium (12)	-0.636		Sedimentation (26)	-0.409
	pH (57)	-0.265			
	Temperature (53)	-0.345	TMINSP	pH (57)	-0.309
	Chloride (55)	-0.337		Sulfate (57)	-0.264
	TDS (55)	-0.343		Ammonia (54)	0.346
	Road Density	-0.221		Chloride (55)	-0.574
	Sedimentation (26)	-0.657		Nitrates (57)	-0.627
				T-Phosphorous (57)	-0.370
PINVER	Magnesium (17)	0.510		Turbidity (48)	0.410

Table 2.9. Continued.

Metric	Variable	r_s	Metric	Variable	r_s
PINVER	Temperature (53)	0.384	TMINSP	TDS (55)	-0.498
	Chloride (55)	-0.511		% Forested	0.380
	Nitrates (57)	-0.591		% Agriculture	-0.263
	O-Phosphates (54)	-0.402		% Urban	-0.208
	T-Phosphorous (57)	-0.422		Road Density	-0.536
	TDS (55)	-0.434		Buffer Road Density	-0.379
	% Forested	0.467			
	% Agriculture	-0.377	TSPECI	Manganese (17)	0.580
	% Urban	-0.359		Nickel (12)	-0.597
	% Buffer Forested	0.354		Sulfate (57)	-0.279
	% Buffer Agriculture	-0.299		Ammonia (54)	0.359
	% Buffer Urban	-0.281		Chloride (55)	-0.518
	Road Density	-0.554		Nitrates (57)	-0.595
	Buffer Road Density	-0.367		T-Phosphorous (57)	-0.281
				Turbidity (48)	0.465
				TSS (55)	0.319
				TDS (55)	-0.559
				% Forested	0.305
				% Agriculture	-0.205
				Road Density	-0.441
				Buffer Road Density	-0.306

Table 2.10. Principal component (PC) loadings of physiochemical data from Ozark Highland stream sites. Each variable contained at least 48 observations. The percent of total model variance accounted for by each component was: PC 1 = 24.2%, PC 2 = 23.3%, PC 3 = 14.0%, PC 4 = 9.9%, and PC 5 = 9.8%.

	Principal Component				
Variable	1	2	3	4	5
DO	0.353	0.167	-0.036	-0.021	0.815
pH	0.147	0.136	0.782	-0.325	-0.071
Temperature	-0.124	-0.262	-0.020	-0.195	-0.805
Sulfate	0.178	0.817	0.018	0.082	0.083
Ammonia	0.175	0.293	-0.813	0.166	-0.172
Chloride	0.347	0.826	0.280	-0.054	0.052
Nitrate	0.417	0.594	0.393	0.090	0.241
O-Phosphorous	0.213	0.865	-0.007	-0.159	0.160
T-Phosphorous	0.240	0.847	0.019	-0.110	0.189
TOC	0.496	0.490	-0.229	0.267	-0.403
Turbidity	-0.125	-0.106	-0.496	0.788	0.128
TSS	0.089	-0.031	-0.284	0.856	0.061
TDS	0.053	0.350	0.773	-0.197	-0.074
% Forested	-0.888	-0.333	-0.027	-0.098	-0.127
% Urban	0.865	0.089	0.145	-0.021	0.061
% Buffer Forested	-0.872	-0.321	0.118	-0.069	-0.133
% Buffer Urban	0.553	0.469	-0.152	-0.311	0.250
Road Density	0.836	0.199	0.163	-0.050	0.162
Buffer Road Density	0.636	0.450	-0.304	-0.270	0.255

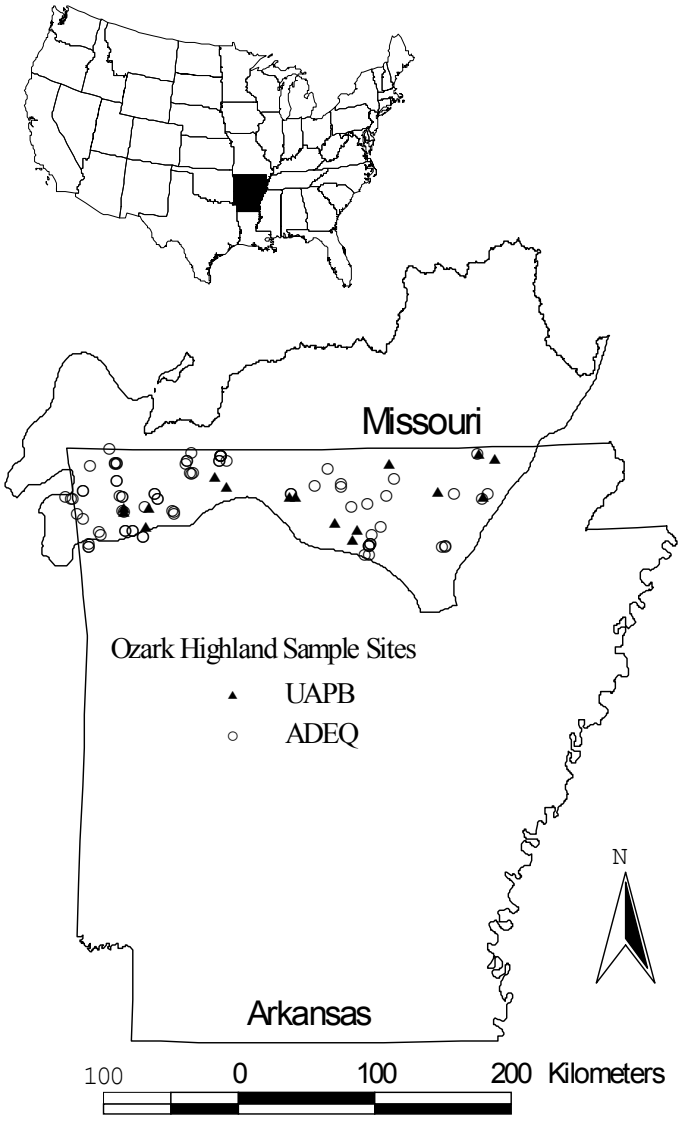


Figure 2.1. Ozark Highland fish-collection sites. Collections were made by the ADEQ (n = 80) or the authors (UAPB; n = 16). All collection sites were listed by the ADEQ as located within the Ozark Highlands ecoregion.

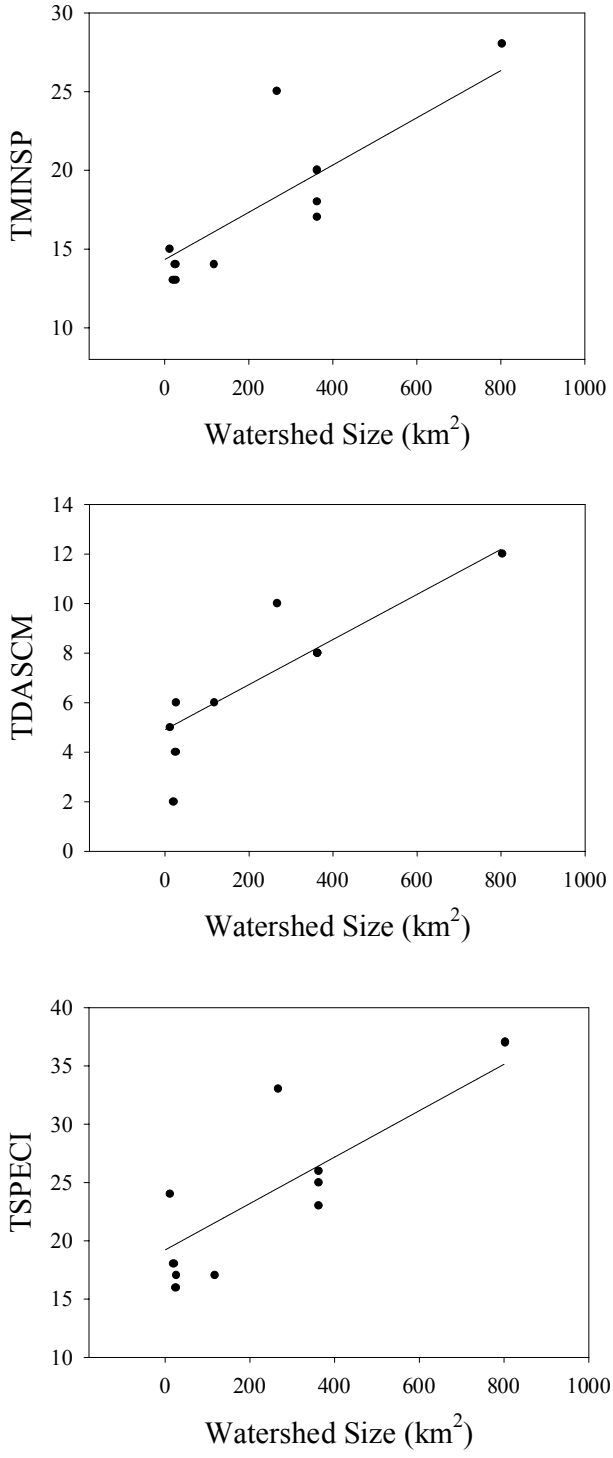


Figure 2.2. Relations between watershed size (km^2) and IBI metrics affected by watershed size (linear regression; $\alpha = 0.05$). Relations were assumed to be linear. Metrics affected by watershed size required linear functions to derive upper threshold limits for metric scoring. n = 10.

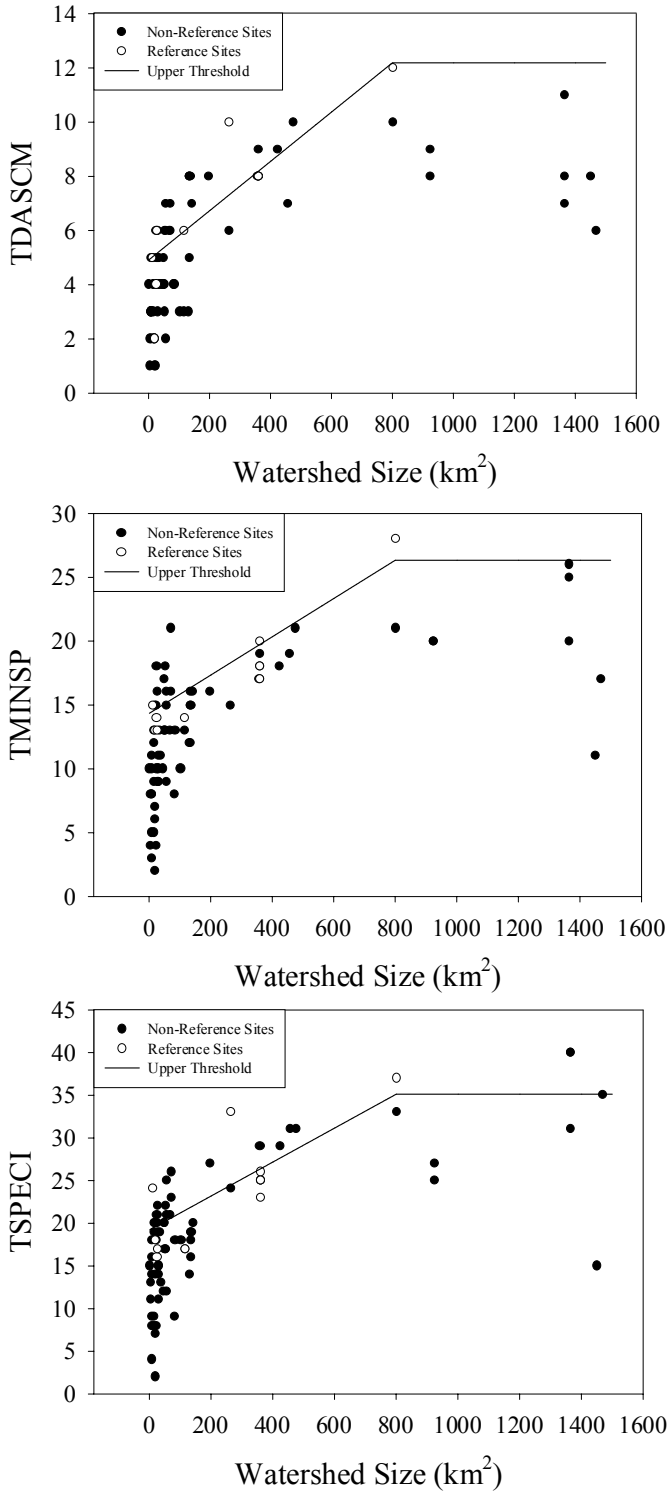


Figure 2.3. Upper threshold limits and raw metric values for TDASCM, TMINSP, and TSPECI used to derive each metric's scoring criteria. Linear functions from 0 – 800 km^2 are best-fit lines for reference-site data. All data n = 77.

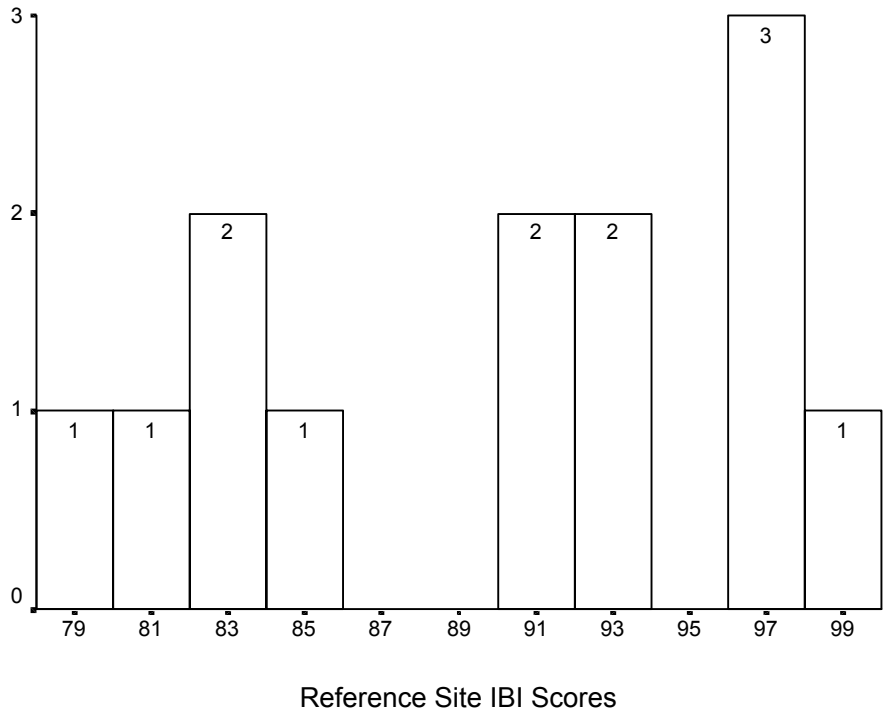


Figure 2.4. Frequency histogram of reference-site IBI scores. Mean = 90.0, SD = 7.0, and n = 13.

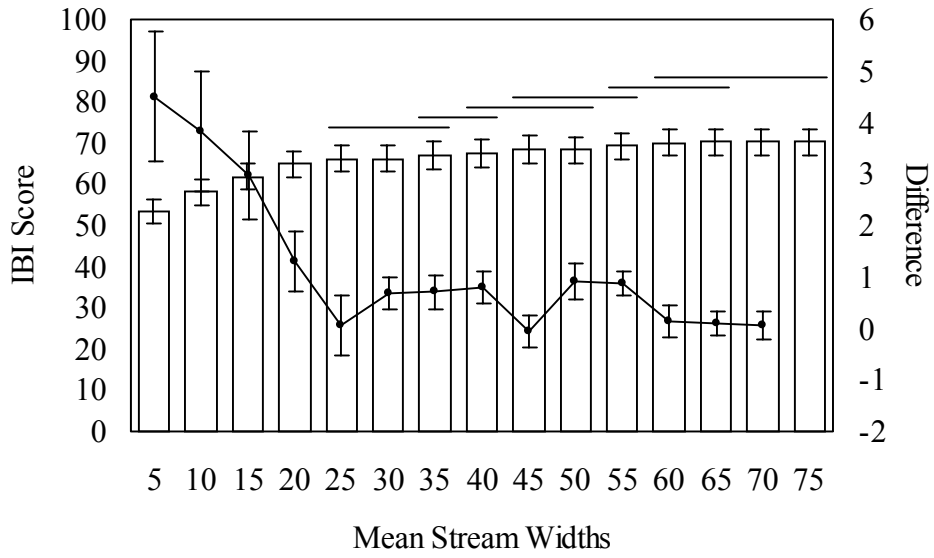


Figure 2.5. Mean IBI scores (± SE; Bars) for every five mean stream widths (MSW) of stream sampled and mean differences in IBI scores (± SE; Points) between adjacent MSW numbers ((n MSWs + 5) – n MSWs) for sample sites (n = 30). Lines overlapping MSWs indicate no significant difference using repeated-measures ANOVA and Tukey test ($\alpha = 0.05$).

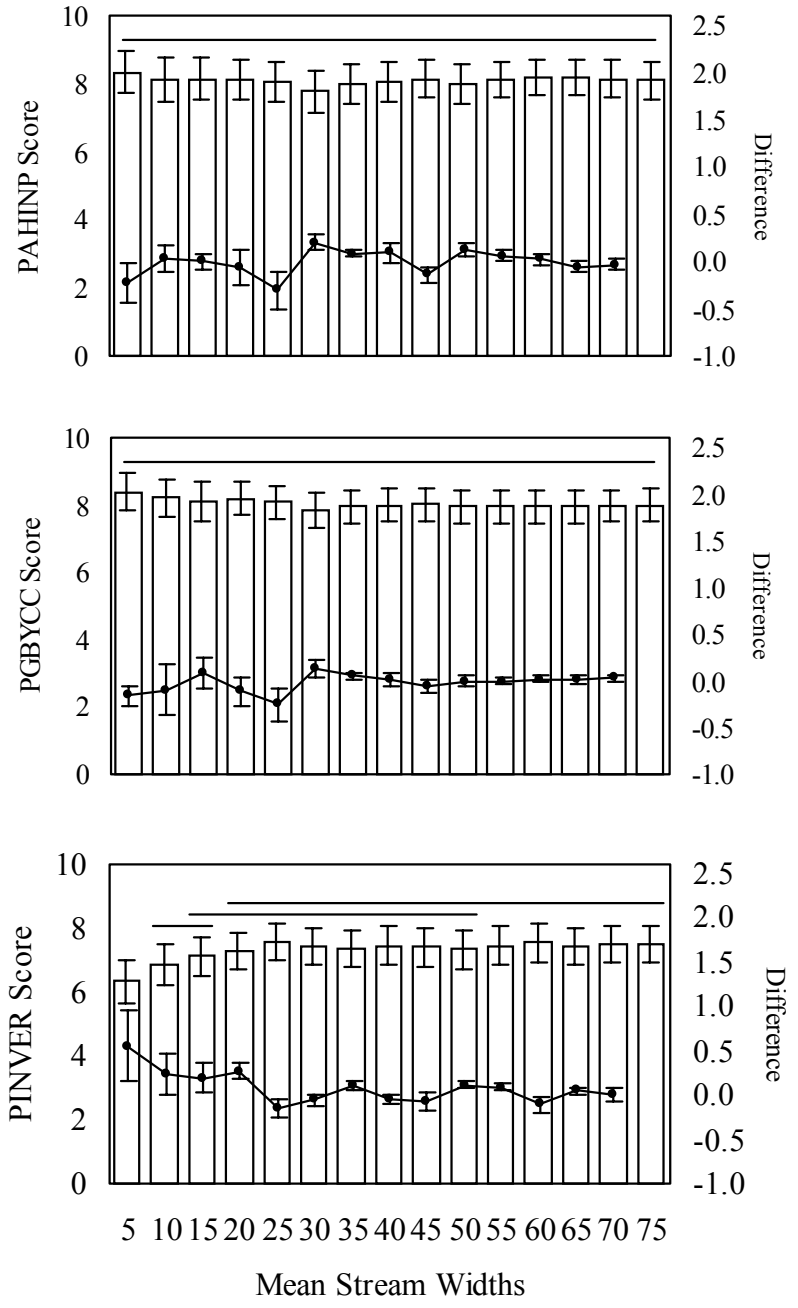


Figure 2.6. Mean metric scores (± SE; Bars) every five mean stream widths (MSW) of stream sampled and mean differences in metric scores (± SE; Points) between adjacent MSW numbers ((n MSWs + 5) – n MSWs) for sample sites (n = 30). Lines overlapping MSWs indicate no significant difference using repeated-measures ANOVA and Tukey test ($\alpha = 0.05$). Metric score scales may change across metrics.

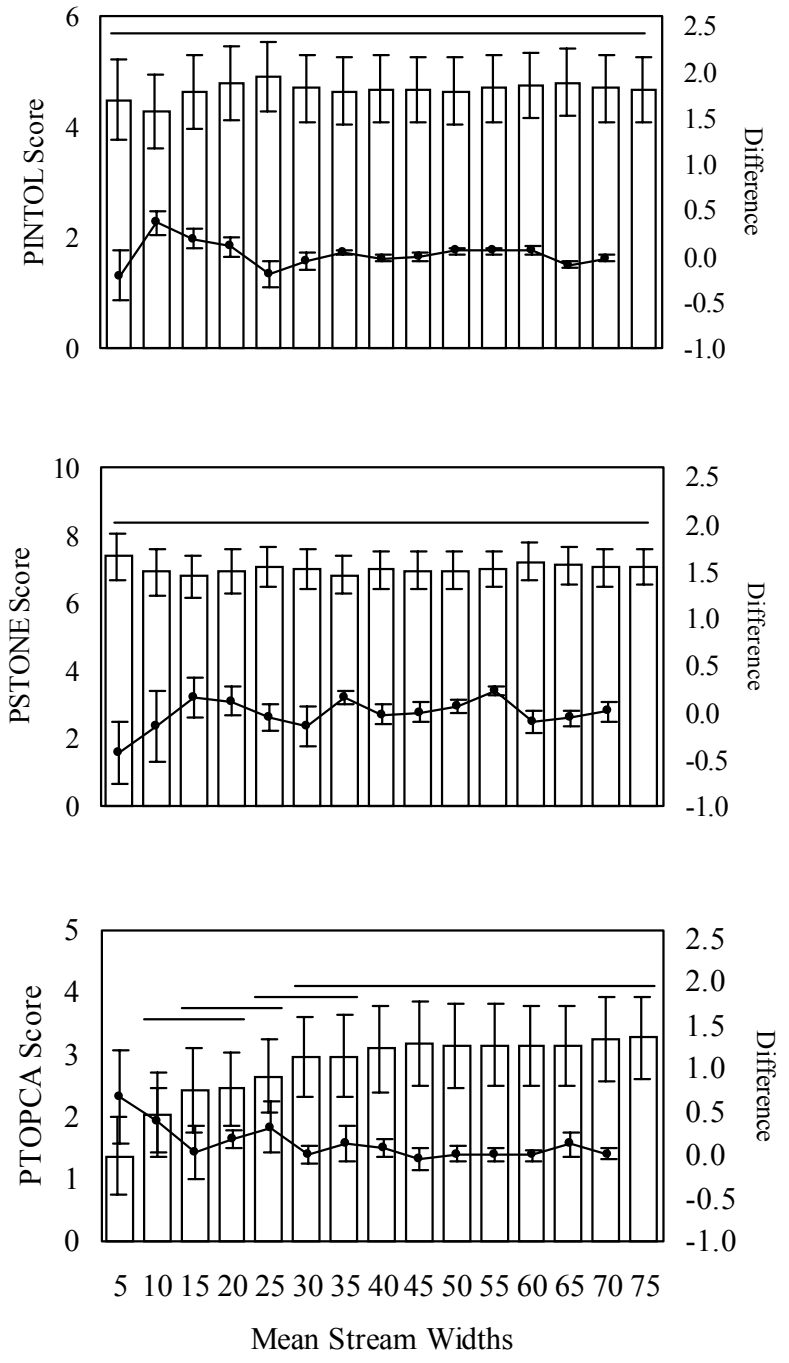


Figure 2.6. Continued.

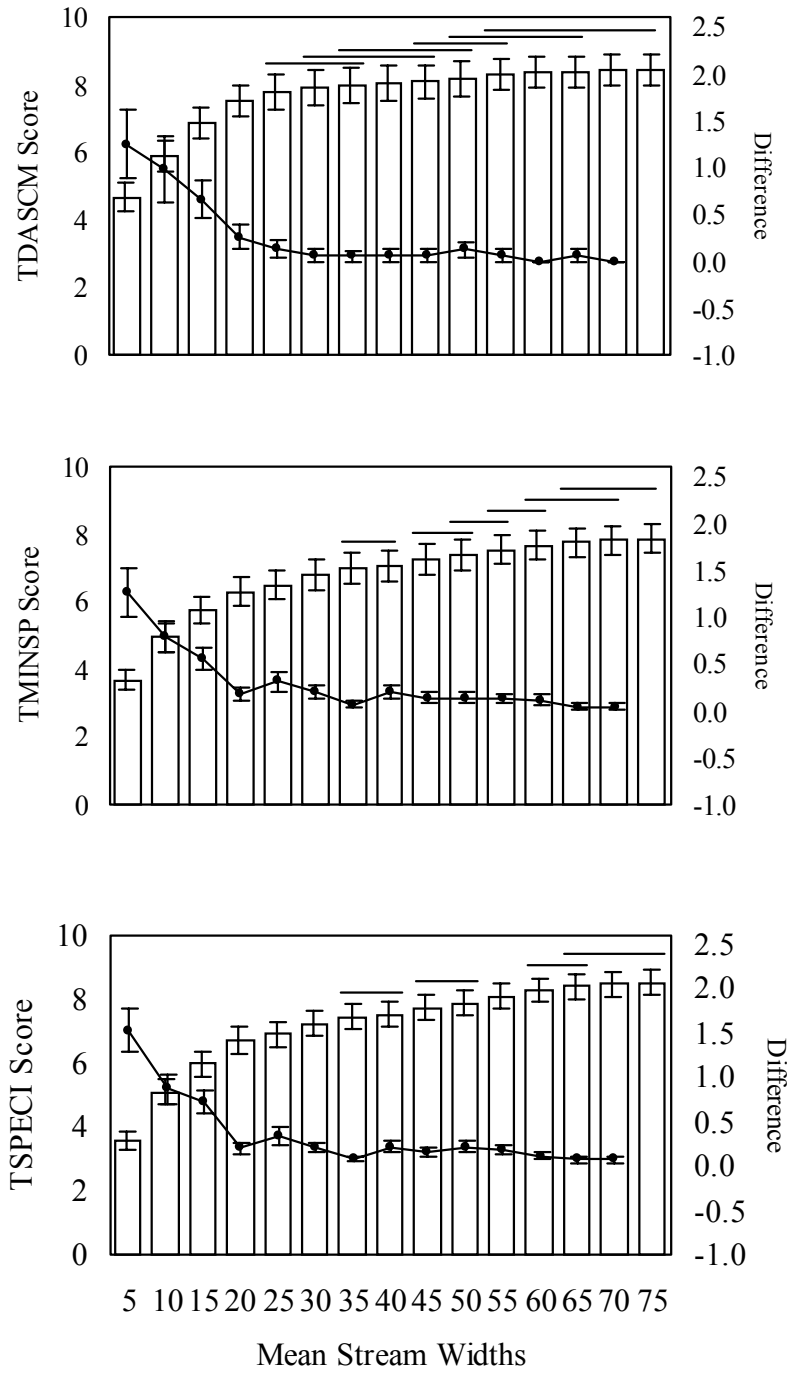


Figure 2.6. Continued.

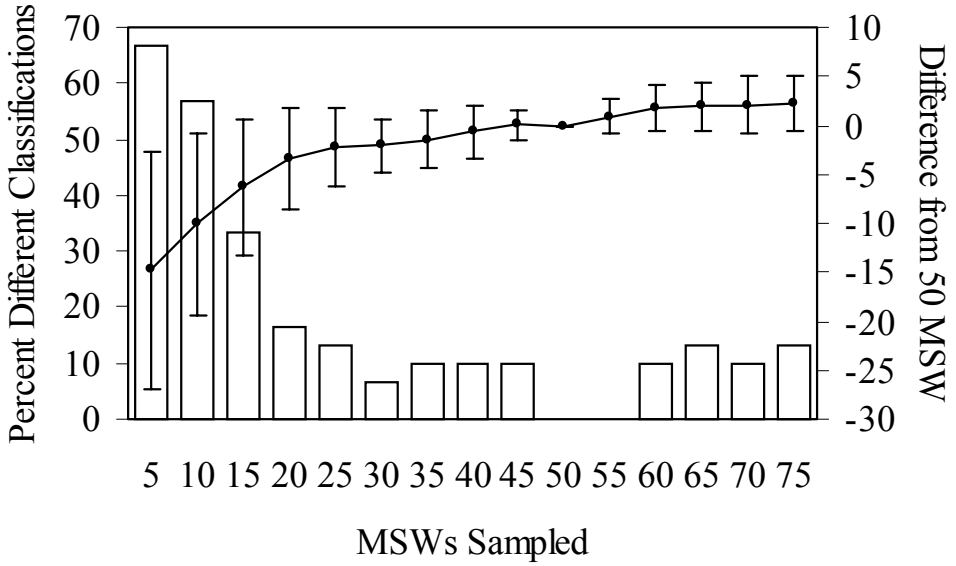


Figure 2.7. The percent of sites having different qualitative site classifications (bars) and mean (± SD; points) difference in IBI scores at each number of MSWs sampled when compared to 50 MSWs. n = 30 for each MSW number.

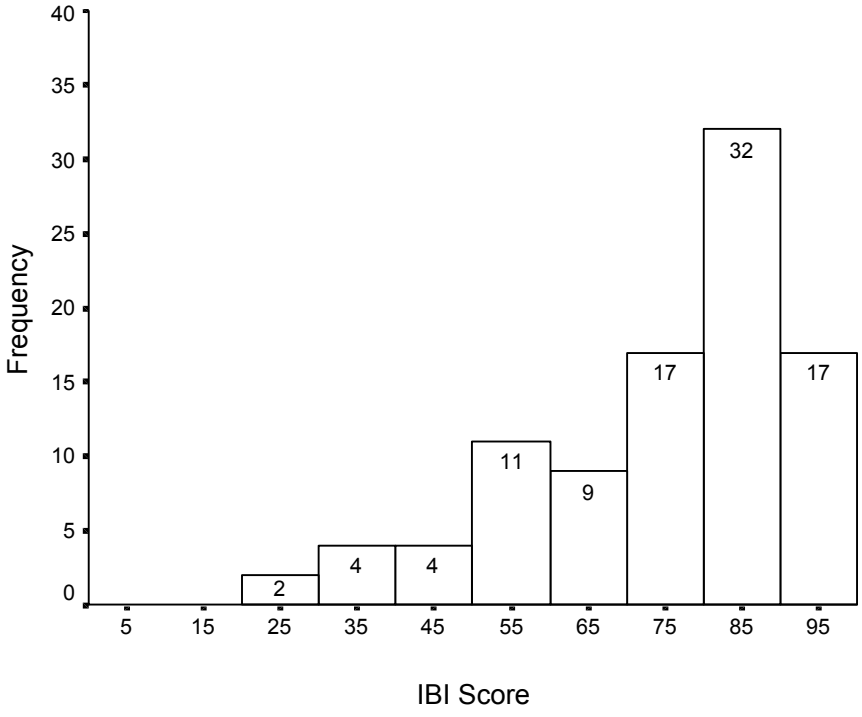


Figure 2.8. Frequency histogram of IBI scores for sites sampled in Arkansas' Ozark Highlands; Mean = 74.8, SD = 17.5, n = 96.

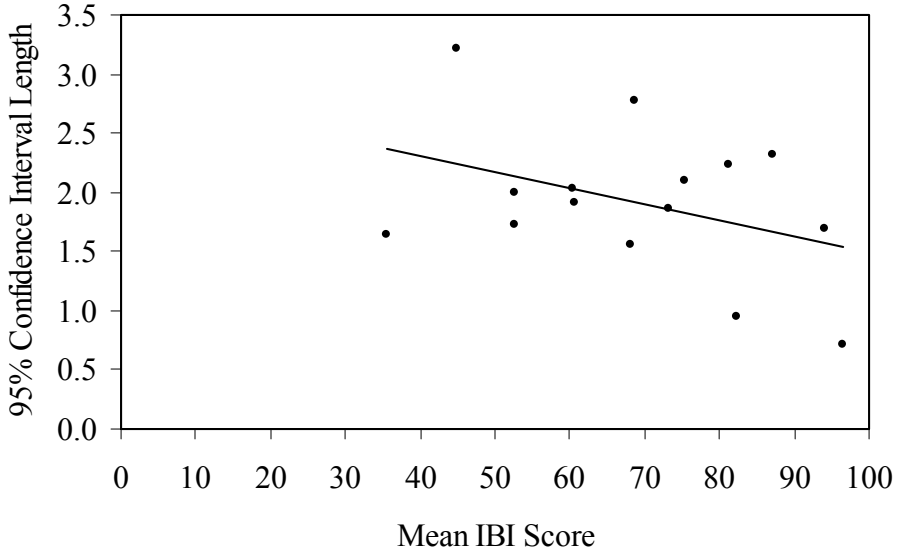


Figure 2.9. The relation between mean IBI scores and 95% confidence interval lengths of IBI scores for 15 Ozark Highland stream sites (including least-squares best-fit line). Index of Biotic Integrity scores for each site were calculated from 25 random samples of 10 (out of 15) stream segments in which fish were collected. Results from linear correlation indicate no significant trend (p-value = 0.48; r = -0.392) between mean IBI scores (i.e., site quality) and 95% confidence interval length (i.e., variability in IBI scores).

Appendix A

Results from the data-rarification process whereby individual fish were randomly eliminated from fish-collection data sets for each site from sampling 75 MSW in order to simulate samples that represented 51 MSW.

Taxon	Times Occurred	Times Eliminated	Percent
Petromyzontidae	8	1	13
Ammocoete	8	1	13
Lepisosteidae	8	3	38
Lepisosteus	8	3	38
Lepisosteus osseus	8	3	38
Clupeidae	12	1	8
Dorosoma	12	1	8
Dorosoma cepedianum	12	1	8
Cyprinidae	714	26	4
Campostoma	99	0	0
Campostoma anomalum	95	0	0
Campostoma oligolepis	4	0	0
Ctenopharyngodon	1	1	100
Ctenopharyngodon idella	1	1	100
Cyprinella	27	1	4
Cyprinella galactura	11	1	9
Cyprinella whipplei	16	0	0
Cyprinus	20	2	10
Cyprinus carpio	20	2	10
Erimystax	10	0	0
Erimystax dissimilis	6	0	0
Erimystax harryi	3	0	0
Erimystax x-punctatus	1	0	0
Hybopsis	14	0	0
Hybopsis amblops	14	0	0
Luxilus	211	3	1
Luxilus cardinalis	19	1	5
Luxilus chrysocephalus	35	2	6
Luxilus pilsbryi	58	0	0
Luxilus zonatus	7	0	0
Lythrurus	4	1	25
Lythrurus umbratilis	4	1	25
Nocomis	46	4	9
Nocomis asper	9	0	0
Nocomis biguttatus	37	4	11
Notemigonus	6	1	17

Taxon	Times Occurred	Times Eliminated	Percent
Notemigonus crysoleucas	6	1	17
Notropis	144	7	5
Notropis boops	37	2	5
Notropis greenei	6	0	0
Notropis nubilus	60	2	3
Notropis ozarcanus	1	0	0
Notropis rubellus	18	0	0
Notropis telescopus	22	3	14
Phoxinus	35	3	9
Phoxinus erythrogaster	35	3	9
Pimephales	53	3	6
Pimephales notatus	48	2	4
Pimephales promelas	3	1	33
Pimephales tenellus	2	0	0
Semotilus	44	0	0
Semotilus atromaculatus	44	0	0
Catostomidae	186	21	11
Catostomus	15	1	7
Catostomus commersoni	15	1	7
Carpiodes	3	0	0
Carpiodes cyprinus	1	0	0
Carpiodes velifer	2	0	0
Erimyzon	11	3	27
Erimyzon oblongus	11	3	27
Hypentelium	70	9	13
Hypentelium nigricans	70	9	13
Ictiobus	2	0	0
Ictiobus bubalus	1	0	0
Ictiobus niger	1	0	0
Minytrema	1	0	0
Minytrema melanops	1	0	0
Moxostoma	84	8	10
Moxostoma carinatum	11	0	0
Moxostoma duquesnei	34	5	15
Moxostoma erythrurum	37	2	5
Moxostoma macrolepidotum	2	1	50
Ictaluridae	183	8	4
Ameiurus	63	6	10
Ameiurus melas	14	4	29
Ameiurus natalis	49	2	4
Ictalurus	10	1	10
Ictalurus punctatus	10	1	10
Noturus	106	1	1
Noturus albater	30	0	0
Noturus exilis	74	1	1

Taxon	Times Occurred	Times Eliminated	Percent
Noturus flavater	2	0	0
Pylodictus	4	0	0
Pylodictus olivaris	4	0	0
Salmonidae	3	1	33
Oncorhyncus	1	1	100
Oncorhyncus mykiss	1	1	100
Salmo	2	0	0
Salmo trutta	2	0	0
Aphredoderidae	3	1	33
Aphredoderus	3	1	33
Aphredoderus sayanus	3	1	33
Fundulidae	101	3	3
Fundulus	101	3	3
Fundulus catenatus	39	1	3
Fundulus notatus	1	0	0
Fundulus olivaceus	61	2	3
Poeciliidae	21	2	10
Gambusia	21	2	10
Gambusia affinis	21	2	10
Atherinidae	12	1	8
Labidesthes	12	1	8
Labidesthes sicculus	12	1	8
Cottidae	71	2	3
Cottus	71	2	3
Cottus carolinae	68	2	3
Cottus hypselurus	3	0	0
Percichthyidae	2	0	0
Morone	2	0	0
Morone chrysops	2	0	0
Centrarchidae	447	35	8
Ambloplites	52	4	8
Ambloplites ariommus	9	2	22
Ambloplites constellatus	40	1	3
Ambloplites rupestris	3	1	33
Lepomis	258	14	5
Lepomis cyanellus	90	4	4
Lepomis gulosus	8	1	13
Lepomis macrochirus	65	6	9
Lepomis megalotis	81	1	1
Lepomis microlophus	6	1	17
Lepomis punctatus	8	1	13
Micropterus	134	16	12
Micropterus dolomieu	54	7	13
Micropterus puntulatus	30	2	7
Micropterus salmoides	50	7	14

Taxon	Times Occurred	Times Eliminated	Percent
Pomoxis	3	1	33
Pomoxis annularis	1	1	100
Pomoxis nigromaculatus	2	0	0
Percidae	381	23	6
Etheostoma	339	18	5
Etheostoma blennioides	55	3	5
Etheostoma caeruleum	63	0	0
Etheostoma euzonum	10	0	0
Etheostoma flabellare	32	2	6
Etheostoma juliae	25	1	4
Etheostoma punctulatum	33	4	12
Etheostoma spectabile	79	3	4
Etheostoma stigmaeum	8	3	38
Etheostoma zonale	34	2	6
Percina	41	5	12
Percina caprodes	36	3	8
Percina evides	1	0	0
Percina nasuta	4	2	50
Stizostedion	1	0	0
Stizostedion vitreum	1	0	0